THESIS / THÈSE

MASTER EN SCIENCES MATHÉMATIQUES

Algorithmes de descente pour la résolution d'inéquations variationnelles

WANT, Caroline

Award date:
1996

Awarding institution:
Universite de Namur

FACULTES UNIVERSITAIRES NOTRE-DAME DE LA PAIX
NAMUR
FACULTE DES SCIENCES

ALGORITHMES DE DESCENTE POUR LA RÉSOLUTION D'INÉQUATIONS VARIATIONNELLES

Promoteur : V.H. NGUYEN

Caroline WANT

Année académique : 1995-1996

Je remercie spécialement Monsieur le Professeur V.H. NGUYEN *pour l'aide et les précieux conseils qu'il m'a apportés tout au long de l'élaboration de ce mémoire.*

Je souhaite également remercier Mademoiselle S. HAUBRUGE *pour le temps qu'elle m'a consacré en répondant à mes diverses questions.*

Je remercie finalement mes parents et mes proches de m'avoir soutenue et supportée tout au long de mes études ainsi que Sophie et Evelyne pour leurs encouragements, leur amitié et leur sourire.

Résumé

Nous présentons dans ce mémoire un cadre algorithmique pour la résolution de problèmes d'inéquation variationnelle non symétriques et de programmes d'optimisation non différentiable. Notre travail est essentiellement basé sur la thèse de M. PATRIKSSON. Cette étude concerne trois algorithmes en particulier : un algorithme d'approximation du coût, un algorithme de descente et un processus itératif défini à partir du principe du problème auxiliaire. Chacun d'eux est basé sur la résolution, à chaque itération, d'un sous-problème qui les différencie. Tout d'abord, nous reformulons de manière équivalente chacun des sous-problèmes sous la forme d'un programme d'optimisation. Nous proposons ensuite une justification de chacune des grandes étapes des processus algorithmiques à travers différents théorèmes. Nous étudions finalement les propriétés de convergence de ces processus itératifs.

Une caractéristique commune aux méthodes employées est que les problèmes auxiliaires qui consistent à trouver une direction de descente sont définis par des fonctions d'approximation qui remplacent la fonction de coût originale.

Abstract

Our work is essentially based on the thesis written by M. PATRIKSSON.

It aims to investigate an algorithm framework for the solution of asymmetric variational inequality problems and non differentiable optimization programs. Our study concerns three algorithms: a cost approximation algorithm, a descent algorithm and an iterative method defined by the auxiliary problem principle. Each of them is based on the solution, at each iteration, of a subproblem. We first give an equivalent formulation of each subproblem in the form of an optimization problem. Then, we propose a theoretical explanation of each step of our algorithmic process. We finally investigate the basic properties of convergence of these iterative methods.

A common feature of the methods studied here is that auxiliary search direction problems are defined by approximating mappings which replace the original cost function in an iterative manner.

Quelques notations

GVIP	Problème généralisé d'inéquation variationnelle.
VIP	Problème d'inéquation variationnelle.
NNLP	Programme non linéaire non différentiable sans contrainte.
CNNLP	Programme non linéaire non différentiable avec contraintes.
AVIP	Problème d'inéquation variationnelle non symétrique.
$\nabla_x f(x,y)$	Gradient de f calculé par rapport à sa première composante x .
$\nabla^2 f(x)$	Matrice hessienne de f en x .
$\xi_f(x)$	Sous-gradient de f en x .
ξ_f^k	Sous-gradient de f en x^k .
ξ_f^*	Sous-gradient de f en x^* .
$\partial f(x)$	Sous-différentiel de f en x .
$\partial^* f(x)$	Gradient généralisé de f en x au sens de Clarke.
$f'(x;d)$	Dérivée directionnelle à droite de f en x dans la direction d .
$f \in C^p$	f est p-fois continûment différentiable.
f^0	Transformée de Fenchel (ou fonction conjuguée) de la fonction f .
m_f	Constante de monotonicité (ou convexité) forte de f .
M_f	Constante de Lipschitz continuité de f .

α_f	Constante de co-coercivité de f .
$\operatorname{Im} f$	Ensemble des images de x par la fonction f .
$\operatorname{dom} f$	Domaine effectif de f .
$\mathcal{L}_X^f$	Ensemble de niveau de f par rapport à l'ensemble X .
$\operatorname{int} X$	Intérieur de l'ensemble X .
$\operatorname{rint} X$	Intérieur relatif de l'ensemble X .
$\underset{x \in X}{\operatorname{argmin}} f(x)$	Ensemble des points de X qui minimisent f .
$\operatorname{aff} X$	Enveloppe affine de l'ensemble X .
2^Y	Ensemble de tous les sous-ensembles de Y .
$N_X(x)$	Cône normal de X en x .
δ_X	Fonction indicatrice de l'ensemble X définie par

$$\delta_X(x) = \begin{cases} 0 & \text{si } x \in X \ , \\ +\infty & \text{sinon.} \end{cases}$$

s.d.p.	Semi-défini positif.
s.c.i.	Semi-continu inférieurement.
$\|d\|$	Norme euclidienne du vecteur d .
$\lambda_{\min}(A)$	Plus petite valeur propre de la matrice A .

Table des Matières

Introduction

Soit X un ensemble convexe, fermé, non vide de $I\!R^n$, et soient $F : X \longrightarrow I\!R^n$ une fonction et $U : X \longrightarrow 2^{I\!R^n}$ une application multivoque monotone.
Nous considérons le problème suivant :

$$\text{(GVIP)} \quad \boxed{\begin{array}{l} \text{Trouver } x^* \in X \text{ et } u^* \in U(x^*) \text{ tels que} \\ \qquad\qquad [F(x^*) + u^*]^T \ (x - x^*) \geq 0 \qquad \forall\, x \in X \ , \end{array}}$$

Ce problème est connu sous le nom de *problème généralisé d'inéquation variationnelle.*

Il possède diverses formulations. En voici deux :

1. Si nous supposons que l'application multivoque U est le sous-différentiel ∂u d'une fonction $u : I\!R^n \longrightarrow I\!R \cup \{+\infty\}$ convexe, propre et semi-continue inférieurement, nous pouvons exprimer (GVIP) sous une forme plus facile à manipuler. Pour ce faire, introduisons une hypothèse de régularité :

   Hypothèse 0.1
   $\text{rint}(\text{dom}\, U) \cap \text{rint}\, X \neq \emptyset$.

   Sous cette hypothèse supplémentaire, nous pouvons montrer, en utilisant [18], [20], [21], et [19, Th. 27.4] que (GVIP) est équivalent au problème suivant :

$$\boxed{\begin{array}{l} \text{Trouver } x^* \in X \text{ tel que} \\ F(x^*)^T(x - x^*) + u(x) - u(x^*) \geq 0 \qquad \forall\, x \in X \ . \end{array}} \tag{0.1}$$

2. Sous cette même hypothèse de régularité (0.1), nous pouvons formuler (GVIP) d'une autre façon équivalente à la première (Annexe A.1) :

$$\boxed{\begin{array}{l}\text{Trouver } x^* \in X \text{ tel que} \\ 0 \in F(x^*) + U(x^*) + N_X(x^*)\,.\end{array}} \tag{0.2}$$

0.1 Cas particuliers de (GVIP)

Dans ce mémoire, nous allons porter attention à deux cas particuliers du problème (GVIP) :

1. les programmes d'optimisation non linéaire, non différentiable sans contrainte (Chapitre 2), caractérisés par :
   - $X = I\!R^n$,
   - $F \equiv \nabla f$ avec $f : I\!R^n \to I\!R \cup \{+\infty\}$ une fonction de classe C^1 sur $\operatorname{dom} u$,
   - $U \equiv \partial u$ avec $u : I\!R^n \to I\!R \cup \{+\infty\}$ une fonction convexe, propre et semi-continue inférieurement.

2. les problèmes d'inéquation variationnelle non symétriques (Chapitre 3) où
   - $U = \partial u$ avec $u : I\!R^n \to I\!R \cup \{+\infty\}$ une fonction convexe, propre et semi-continue inférieurement,
   - F et X identiques à ceux définis pour (GVIP).

Nous attirons à présent votre attention sur le fait que toutes les fonctions avec lesquelles nous allons travailler sont convexes. Ce mémoire fera donc appel à de nombreux arguments d'analyse convexe. C'est pourquoi nous rappellerons, au Chapitre 1, toute une série de résultats importants dont nous aurons besoin dans la suite, la plupart ayant trait à l'analyse convexe.

0.2 Les algorithmes de résolution des cas particuliers de (GVIP)

Pour chacune de ces particularisations, nous avons pour objectif de développer un cadre algorithmique permettant de déterminer leur(s) solution(s) et d'étudier les propriétés de convergence de ces processus itératifs.

- Pour les programmes d'optimisation non différentiable, nous étudierons deux algorithmes. Nous développerons tout d'abord l'algorithme d'approximation du coût dû à M. Patriksson. Une itération de ce dernier est divisée en deux étapes que nous pouvons brièvement décrire comme suit :

  1. Étant donné un point admissible x , nous définissons une direction de recherche $d = y - x$ où y est la solution (parfois inexacte) d'un problème approximé du problème initial. Nous entendons par problème approximé un problème où la fonction F est approximée par un opérateur quelconque.

  2. La direction de recherche obtenue par le point 1. est une direction de descente pour une fonction de mérite ψ dont les minima coïncident avec l'ensemble des solutions de (GVIP). Ainsi, nous pouvons effectuer une recherche linéaire exacte relativement à cette fonction de mérite dans la direction de descente obtenue en 1. Cette application multivoque algorithmique de recherche linéaire définit un nouvel itéré qui réduit la valeur de la fonction de mérite.

  Nous exposerons ensuite le deuxième algorithme qui est appelé *processus itératif de Cohen*. Ce dernier diffère de l'algorithme d'approximation du coût par le fait qu'aucune recherche linéaire n'y est effectuée. On verra cependant que leur formulation est équivalente. Chaque itération du processus itératif de Cohen est basée sur la résolution du problème auxiliaire de Cohen. Sous sa forme générale, celui-ci s'exprime comme :

  Soient $\varepsilon > 0$, $x \in X$ et Φ une fonction approximant F .
  Trouver $x^+ \in X$ tel que $\forall\, y \in X$

$$[\varepsilon F(x) + \Phi(x^+) - \Phi(x)]^T (y - x^+) + \varepsilon(u(y) - u(x^+)) \geq 0 \,.$$

- En ce qui concerne la résolution des problèmes d'inéquation variationnelle non symétriques, nous exposerons trois algorithmes. Le premier d'entre eux est celui d'approximation du coût développé dans un cas restrictif puisque nous choisirons une longueur de pas égale à l'unité. Ceci est indispensable car aucune fonction de mérite permettant la recherche linéaire n'est connue.

  Le second processus consiste en une variante de l'algorithme du coût d'approximation. Nous l'appellerons *algorithme de descente* car sa particularité réside dans la recherche d'une direction de descente. Une itération de cet algorithme est également divisée en deux parties :

  1. construction d'une fonction $L(x,y)$ dont nous calculons le supremum par rapport à y pour une valeur de x fixée. À partir de la variable y donnant le supremum de $L(x,y)$, nous définissons la direction de descente $d = y - x$.
  2. En posant
  $$\psi(x) = \sup_{y \in X} L(x,y) ,$$
  nous définissons une fonction de mérite ψ . Nous pouvons dès lors effectuer une recherche linéaire exacte le long de la direction d , obtenue en 1., relativement à ψ .

  Le troisième algorithme que nous traiterons est celui de Cohen déjà cité ci-dessus.

0.3 Convergence des processus itératifs

Comme nous l'avons déjà mentionné, un de nos objectifs principaux est d'étudier les conditions de convergence de chacun des algorithmes développés pour résoudre les cas particuliers de (GVIP). Nous porterons une grande attention aux hypothèses de convergence de ces algorithmes et principalement celles du processus itératif de Cohen. En effet, le quatrième chapitre du mémoire sera consacré à l'extension de la région de convergence de l'algorithme de Cohen. Nous savons déjà qu'il converge sous l'hypothèse de forte monotonicité de la fonction F . Nous introduirons, dans le Chapitre 4, la notion de co-coercivité d'une fonction et nous prouverons la convergence du processus itératif de Cohen sous l'hypothèse moins forte de co-coercivité.

Afin d'élaborer ce mémoire, nous nous sommes principalement basés sur les ouvrages suivants :

- M. PATRIKSSON, A unified framework of descente algorithms for nonlinear programs and variational inequalities, Division of Optimization, Department of Mathematics, Linköping University, Sweden, 1993.

- M. PATRIKSSON, A descent algorithm for a class of variational inequalities, preprint, University of Washington, March 31, 1995.

- D.L. ZHU, & P. MARCOTTE, Co-coercivity and its role in the convergence of iterative schemes for solving variational inequalities, preprint, Université de Montréal, 1995.

- A. RENAUD, & G. COHEN, An extension of the auxiliary problem principle to non-symmetric auxiliary operators, preprint, École des Mines de Paris, August 22, 1995.

Chapitre 1

Préliminaires

Dans ce chapitre, nous rassemblons différentes définitions et propriétés intéressantes dont nous aurons fréquemment besoin dans le mémoire. Il s'agit tout d'abord de diverses propriétés de fonctions telles que la différentiabilité, la convexité, la monotonicité, la semi-continuité inférieure, la fermeture et la coercivité. Ensuite, nous exposons des définitions et des critères se rapportant à l'optimalité. Nous définissons entre autres la notion de sous-gradient, de direction de descente, de sous-différentiel, de dérivée directionnelle et de cône normal. Nous énonçons une caractérisation des minima des fonctions convexes non nécessairement différentiables qui sera évoquée à maintes reprises tout au long du mémoire. Pour terminer, nous rappelons le théorème A de Zangwill, le concept d'intérieur relatif d'un ensemble, celui de transformée de Fenchel d'une fonction ainsi que les définitions des applications multivoques de recherche d'une direction de descente et de recherche linéaire exacte.

1.1 Diverses propriétés d'une fonction

1.1.1 Différentiabilité

Définition 1.1

Soit X un ensemble non vide de $I\!R^n$ et soit $f : X \to I\!R$ une fonction.

f est *différentiable en* $x_0 \in \operatorname{int} X$ si il existe un vecteur $\nabla f(x_0)$ appelé *gradient de* f *en* x_0 et une fonction $\alpha : I\!R^n \to I\!R$ tels que

$$f(x) = f(x_0) + \nabla f(x_0)^T(x - x_0) + \|x - x_0\| \; \alpha(x_0; x - x_0) \qquad \forall\, x \in X$$

avec $\lim\limits_{x \to x_0} \alpha(x_0; x - x_0) = 0$.

Définition 1.2
f est dite *différentiable sur* $Y \subseteq X$ si et seulement si elle est différentiable en tout point x_0 de Y .

Définition 1.3
Lorsque f est différentiable en x_0 , il existe un unique gradient $\nabla f(x_0)$ défini par

$$\nabla f(x_0) = \left(\frac{\partial f(x_0)}{\partial x_1}, \cdots, \frac{\partial f(x_0)}{\partial x_n}\right)^T$$

où $\frac{\partial f(x_0)}{\partial x_i}$ est la dérivée partielle de f par rapport à x_i en x_0 .

Définition 1.4
Soit X un ensemble non vide de $I\!R^n$ et soit $f : X \to I\!R$.
f est *deux fois différentiable en* $x_0 \in \operatorname{int} X$ si il existe un vecteur $\nabla f(x_0)$ et une matrice symétrique $n \times n$, $\nabla^2 f(x_0)$, appelée *matrice hessienne* et une fonction $\alpha : I\!R^n \to I\!R$ tels que

$$\begin{aligned} f(x) &= f(x_0) + \nabla f(x_0)^T(x - x_0) + \frac{1}{2}(x - x_0)^T \nabla^2 f(x_0)(x - x_0) \\ &\quad + \|x - x_0\|^2 \alpha(x_0; x - x_0) \qquad \forall x \in X \end{aligned}$$

avec $\lim_{x \to x_0} \alpha(x_0; x - x_0) = 0$.

Définition 1.5
f est dite *deux fois différentiable sur* $Y \subseteq X$ si et seulement si elle est deux fois différentiable en tout point x_0 de Y .

Définition 1.6
Lorsque f est deux fois différentiable, la matrice hessienne $\nabla^2 f(x_0)$ est définie par

$$\nabla^2 f(x_0) = \begin{pmatrix} \frac{\partial^2 f(x_0)}{\partial x_1^2} & \frac{\partial^2 f(x_0)}{\partial x_1 x_2} & \cdots & \frac{\partial^2 f(x_0)}{\partial x_1 x_n} \\ \frac{\partial^2 f(x_0)}{\partial x_2 x_1} & \frac{\partial^2 f(x_0)}{\partial x_2^2} & \cdots & \frac{\partial^2 f(x_0)}{\partial x_2 x_n} \\ \vdots & \vdots & \vdots & \vdots \\ \frac{\partial^2 f(x_0)}{\partial x_n x_1} & \frac{\partial^2 f(x_0)}{\partial x_n x_2} & \cdots & \frac{\partial^2 f(x_0)}{\partial x_n^2} \end{pmatrix} .$$

1.1.2 Convexité

Soit X un ensemble convexe fermé non vide de $I\!R^n$ et soit f une fonction définie de X dans $I\!R$.

Définition 1.7
f est *convexe sur* X si et seulement si

$$\forall\ x, y \in X\ , \quad \forall\ \lambda \in [0,1]$$
$$f(\lambda x + (1-\lambda)y) \leq \lambda\ f(x) + (1-\lambda)\ f(y)\ .$$

Proposition 1.1
Si $f \in C^1$ *sur* X *,*
alors f *est* convexe sur X *si et seulement si*

$$f(x) \geq f(y) + \nabla f(y)^T\ (x-y) \qquad \forall\ x, y \in X$$

ou de manière équivalente si et seulement si

$$[\nabla f(x) - \nabla f(y)]^T\ (x-y) \geq 0 \qquad \forall\ x, y \in X$$

cfr. [2, Th. 3.3.3 et Th. 3.3.4].

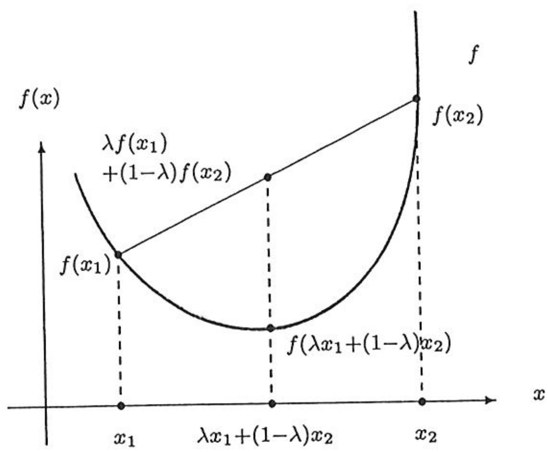

Figure 1.1. Interprétation géométrique de la convexité d'une fonction.

Définition 1.8
f est *strictement convexe sur* X si et seulement si,

$$\forall\ x,y \in X\ , \forall\ \lambda \in [0,1]$$
$$f(\lambda x + (1-\lambda)y) < \lambda f(x) + (1-\lambda) f(y)\ .$$

Proposition 1.2
Si $f \in C^1$ *sur* X ,
alors f *est* strictement convexe sur X *si et seulement si*

$$f(x) > f(y) + \nabla f(y)^T\ (x-y) \qquad \forall\ x,y \in X\ ,\ x \neq y$$

ou de manière équivalente si et seulement si

$$[\nabla f(x) - \nabla f(y)]^T\ (x-y) > 0 \qquad \forall\ x,y \in X\ ,\ x \neq y$$

cfr. [2, Th. 3.3.3 et Th.3.3.4].

Définition 1.9
f est *fortement convexe de module* m_f *sur* X si et seulement si $\exists\ m_f > 0$ tel que

$$\forall\ x,y \in X\ ,\ \ \forall\ \lambda \in [0,1]$$

$$f(\lambda x + (1-\lambda)y) \leq \lambda f(x) + (1-\lambda) f(y) - \frac{m_f}{2} \lambda(1-\lambda)\ \|x-y\|^2\ .$$

Proposition 1.3
Si $f \in C^1$ *sur* X ,
alors f *est* fortement convexe de module m_f sur X *si et seulement si* $\exists\ m_f > 0$ *tel que*

$$f(x) \geq f(y) + \nabla f(y)^T\ (x-y) + \frac{m_f}{2}\ \|x-y\|^2 \qquad \forall\ x,y \in X$$

ou de manière équivalente si et seulement si

$$[\nabla f(x) - \nabla f(y)]^T\ (x-y) \geq m_f\ \|x-y\|^2 \qquad \forall\ x,y \in X$$

cfr. [17, Sec. 1.1.4, Le 3] et [9].

Proposition 1.4
Si $f \in C^2$ *sur* X ,
alors f *est* fortement convexe de module m_f sur X *si et seulement si* $\exists\ m_f > 0$ *tel que*

$$y^T(\nabla^2 f(x) - m_f\ I)\ y \geq 0 \qquad \forall\ x \in X\ ,\ \forall\ y \in I\!R^n$$

cfr. [17, Sec. 1.1.4].

1.1.3 Monotonicité

Soient X un ensemble convexe fermé non vide de $I\!R^n$ et f une fonction de X dans $I\!R^n$.

Définition 1.10
f est *monotone sur* X si et seulement si

$$[f(x) - f(y)]^T(x-y) \geq 0 \qquad \forall\, x, y \in X\ .$$

Proposition 1.5
Si $f \in C^1$ *sur* X *,*
alors f *est* monotone sur X *si et seulement si*

$$y^T\ \nabla f(x)\ y \geq 0 \qquad \forall\, x \in X\ ,\ \forall\, y \in I\!R^n$$

cfr. [14, Th. 5.4.3].

Définition 1.11
f est *strictement monotone sur* X si et seulement si

$$[f(x) - f(y)]^T\ (x-y) > 0 \qquad \forall\, x, y \in X\ ,\ x \neq y\ .$$

Définition 1.12
f est *fortement monotone de module* m_f *sur* X si et seulement si $\exists\ m_f > 0$ tel que

$$[f(x) - f(y)]^T\ (x-y) \geq m_f\ \|x-y\|^2 \qquad \forall\, x, y \in X\ .$$

Proposition 1.6
Si $f \in C^1$ *sur* X *,*
alors f *est* fortement monotone de module m_f sur X *si et seulement si* $\exists\ m_f > 0$ *tel que*

$$y^T\ (\nabla f(x) - m_f\ I)\ y \geq 0 \qquad \forall\, x \in X\ ,\ \forall\, y \in I\!R^n$$

cfr. [14, Th. 5.4.3].

1.1.4 Continuité lipschitzienne

Définition 1.13
Soit X un ensemble convexe fermé non vide de $\mathbb{R}^n$ et soit $f : X \to \mathbb{R}^n$.
f est une *fonction Lipschitz continue de module M_f sur X* si et seulement si $\exists\, M_f \geq 0$ tel que

$$\|f(x) - f(y)\| \leq M_f \; \|x - y\| \qquad \forall\, x, y \in X \; .$$

Proposition 1.7
Soit $f \in C^1$ sur X .
Si ∇f est Lipschitz continu de module $M_{\nabla f}$ sur X *, alors il existe $M_{\nabla f} \geq 0$ tel que*

$$f(y) - f(x) \leq \nabla f(x)^T \; (y - x) + \frac{M_{\nabla f}}{2} \; \|x - y\|^2 \qquad \forall\, x, y \in X \; .$$

La démonstration de cette propriété est réalisée en annexe A.2.

1.1.5 Semi-continuité inférieure

Soit une fonction $f : \mathbb{R}^n \to \mathbb{R} \cup \{-\infty, +\infty\}$.

Définition 1.14
f est *semi-continue inférieurement (s.c.i.) en $x_0 \in \mathbb{R}^n$* si et seulement si $\forall\, \varepsilon > 0 \quad \exists\, \delta > 0$ tel que

$$\|x - x_0\| \leq \delta \quad \Rightarrow \quad f(x) > f(x_0) - \varepsilon \qquad \forall\, x \in \mathbb{R}^n \; .$$

Définition 1.15
f est *s.c.i. sur $\mathbb{R}^n$* si et seulement si elle l'est pour tout point $x_0 \in \mathbb{R}^n$.

Proposition 1.8
f est s.c.i. sur $\mathbb{R}^n$ *si et seulement si*

$$f(x) = \liminf_{y \to x} f(y) \qquad \forall\, x \in \mathbb{R}^n$$

cfr. [19].

Proposition 1.9
Toute fonction continue est s.c.i.
La réciproque n'est cependant pas toujours vraie.

1.1.6 Fonction propre

Soit une fonction $f : \mathbb{R}^n \to \mathbb{R} \cup \{-\infty, +\infty\}$.

Définition 1.16
On appelle *domaine effectif de* f l'ensemble noté $\operatorname{dom} f$ et défini par

$$\operatorname{dom} f = \{x \in \mathbb{R}^n \mid f(x) < +\infty\} .$$

Définition 1.17
Une fonction est *propre* si et seulement si $\operatorname{dom} f \neq \emptyset$ et $f(x) > -\infty$ pour tout $x \in \mathbb{R}^n$.

1.1.7 Coercivité

Soient X un ensemble convexe fermé non vide de $\mathbb{R}^n$ et f une fonction de X dans $\mathbb{R}$.

Définition 1.18
f est *faiblement coercive sur* X si et seulement si

$$\lim_{\substack{x \in X \\ \|x\| \to +\infty}} f(x) = +\infty .$$

Définition 1.19
f est *coercive sur* X si et seulement si

$$\lim_{\substack{x \in X \\ \|x\| \to +\infty}} \frac{f(x)}{\|x\|} = +\infty .$$

1.1.8 Fermeture

Soient X et Y deux sous-espaces de $\mathbb{R}^n$ et $f : X \to 2^Y$ une application multivoque.

Définition 1.20
f est *fermée en* $x \in X$ si et seulement si

$$\left.\begin{array}{l} \{x^k\} \to x \\ y^k \in f(x^k) \quad \forall\, k \\ \{y^k\} \to y \end{array}\right\} \quad \Rightarrow \quad y \in f(x)$$

cfr. [13], [8].

Définition 1.21
Si f est fermée en tout point $x \in X$, alors f est *fermée sur* X.

1.2 Optimalité

1.2.1 Admissibilité

Définition 1.22
Soit (P) un problème d'optimisation avec contraintes. On appelle *point admissible pour le problème* (P) tout point qui vérifie les contraintes de (P) .

Définition 1.23
Soit $d \in I\!R^n$, $d \neq 0$ est une *direction admissible pour* X *à partir de* $x_0 \in X$ si et seulement si

$$\exists\ \delta > 0 \text{ tel que } \quad \forall\ t \in [0, \delta\,[\qquad x_0 + td \in X\ .$$

1.2.2 Dérivée directionnelle

Soient $f : I\!R^n \to I\!R \cup \{-\infty, +\infty\}$, $x_0 \in I\!R^n$, $f(x_0) < +\infty$, $d \in I\!R^n$, $d \neq 0$.

Définition 1.24
La *dérivée directionnelle à droite de* f *en* x_0 *dans la direction* d est définie par

$$f'(x_0; d) = \lim_{\substack{t \to 0 \\ >}} \frac{f(x_0 + td) - f(x_0)}{t}\ .$$

Proposition 1.10
Si f *est une fonction convexe sur* $I\!R^n$ *et* $x_0 \in \operatorname{dom} f$,
alors $f'(x_0; d)$ *existe pour tout* $d \in I\!R^n$,
cfr. [2, Th. 3.1.5].

1.2.3 Sous-gradient et sous-différentiel

Soient $f : I\!R^n \to I\!R \cup \{-\infty, +\infty\}$ une fonction convexe, $x_0 \in \operatorname{dom} f$ et $y \in I\!R^n$.

Définition 1.25
On appelle *sous-gradient de f en x_0* tout $y \in I\!R^n$ tel que

$$y^T(x - x_0) \leq f(x) - f(x_0) \qquad \forall\, x \in I\!R^n .$$

Définition 1.26
L'ensemble $\partial f(x_0)$ de tels sous-gradients est appelé le *sous-différentiel de f en x_0* .

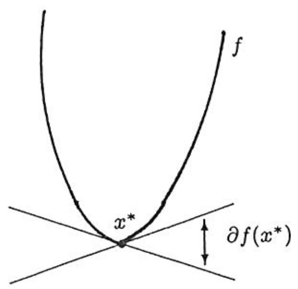

Figure 1.2. Interprétation géométrique du sous-différentiel.

Proposition 1.11

$$y \in \partial f(x_0) \;\Leftrightarrow\; y^T d \leq f'(x_0; d) \qquad \forall\, d \in I\!R^n$$

cfr. [19, Th. 23.2].

Proposition 1.12
Si f est une fonction convexe et propre, alors pour tout $x_0 \in \operatorname{rint}(\operatorname{dom} f)$, $\partial f(x_0)$ est non vide et

$$f'(x_0; d) = \sup_{y \in \partial f(x_0)} y^T d \qquad \forall\, d \in I\!R^n$$

cfr. [19, Th. 23.4].

Proposition 1.13
Si $f \in C^1$, *alors* $\partial f(x_0) = \nabla f(x_0)$ *et*

$$f'(x_0; d) = \nabla f(x_0)^T d$$

cfr. [19, pp. 213 et 216].

1.2.4 Direction de descente

Soient $f : \mathbb{R}^n \to \mathbb{R} \cup \{-\infty, +\infty\}$ une fonction convexe et $d \in \mathbb{R}^n$, $d \neq 0$, $x_0 \in \operatorname{dom} f$.

Définition 1.27
d est une *direction de descente pour* f *en* x_0 si et seulement si $\exists\ \delta > 0$ tel que

$$\forall\ t \in\ ]0, \delta\ [\qquad f(x_0 + td) < f(x_0)\ .$$

Proposition 1.14
Si $f \in C^1$,
alors d *est une* direction de descente pour f en x_0 *si et seulement si*

$$\nabla f(x_0)^T d < 0\ .$$

Proposition 1.15
Si $f \notin C^1$,
alors $d \in \mathbb{R}^n$, $d \neq 0$, *est une* direction de descente pour f en x_0 *si et seulement si* $f'(x_0; d) < 0$
ou de manière équivalente si et seulement si

$$y^T d < 0 \qquad \forall\ y \in \partial f(x_0)$$

ou encore si et seulement si

$$0 \notin \partial f(x_0)\ .$$

1.2.5 Cône normal

Définition 1.28
Le *cône normal à un ensemble convexe* X *de* $\mathbb{R}^n$ *au point* $x_0 \in X$ est noté $N_X(x_0)$ et défini par

$$N_X(x_0) = \{v \in \mathbb{R}^n \text{ tel que } v^T(x - x_0) \leq 0 \quad \forall\, x \in X\}\,.$$

Proposition 1.16
Le cône normal d'un ensemble convexe X en x_0 *peut encore être défini (cfr. [19]) comme* $N_X(x_0) = \partial\delta_X(x_0)$ *où* δ_X *est la fonction indicatrice de l'ensemble* X *:*

$$\delta_X(x_0) = \begin{cases} 0 & si\ x_0 \in X\,, \\ +\infty & si\ x_0 \notin X\,. \end{cases}$$

Soit X un ensemble de contraintes, $x^* \in X$.

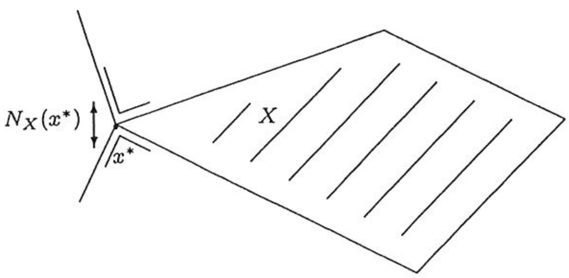

Figure 1.3. Interprétation géométrique du cône normal de X en x^* .

1.2.6 Caractérisation des extrema

Proposition 1.17
Soient $f : \mathbb{R}^n \to \mathbb{R} \cup \{-\infty, +\infty\}$ *une fonction convexe et propre,* $x_0 \in \operatorname{dom} f$.
f *atteint sa valeur minimale en* x_0 *si et seulement si* $0 \in \partial f(x_0)$,
cfr. [12, Th. 4.36].

1.2.7 Conditions d'unicité d'un minimum

Proposition 1.18
Soit X un ensemble convexe de $I\!R^n$.
Si $f : X \to I\!R$ est une fonction strictement convexe, alors f admet au plus un minimum, cfr. [12, Cor. 4.35].

Proposition 1.19
Soit $f : X \to I\!R \cup \{-\infty, +\infty\}$ une fonction fortement convexe de constante m_f , propre et semi-continue inférieurement.
Alors f possède un unique minimum x_0 et celui-ci vérifie

$$f(x_0) \leq f(x) - \frac{m_f}{2}\|x - x_0\|^2 \qquad \forall\, x \in X .$$

La démonstration de ceci est effectuée en annexe A.3.

1.3 Théorème A de Zangwill

Nous rappelons le célèbre théorème A de Zangwill car celui-ci sera utilisé à plusieurs reprises pour démontrer la convergence des algorithmes que nous allons développer pour résoudre les inéquations variationnelles. Ce théorème est tiré de [26].

Soient un problème d'optimisation sur X et Ω l'ensemble des points satisfaisant une certaine contrainte nécessaire d'optimalité. Soit $A : X^n \to X$ une application multivoque (un algorithme par exemple) et considérons une suite $\{x^k\}$ engendrée par l'algorithme, c'est-à-dire vérifiant $x^{k+1} \in A(x^k) \quad \forall\, k$.
Si les trois conditions suivantes sont satisfaites :

1. *les points x^k sont tous contenus dans un compact $C \subseteq X$.*
2. *il existe une fonction de descente T telle que*
   - *si $x^k \in X - \Omega$, alors $\forall\, x^{k+1} \in A(x^k)$*
   $$T(x^{k+1}) < T(x^k) .$$
   - *si $x^k \in \Omega$, alors l'algorithme se termine.*
3. *l'application A est fermée sur $X - \Omega$ et $\forall\, x \in X - \Omega \quad A(x) \neq \emptyset$,*

alors tout point d'accumulation de la suite $\{x^k\}$ appartient à Ω .

1.4 Intérieur relatif d'un ensemble

Définition 1.29
Soit X un ensemble de $\mathbb{R}^n$.
On appelle *enveloppe affine de* X l'ensemble noté $\operatorname{aff} X$ et défini comme le plus petit ensemble affine contenant X .

Définition 1.30
Soit X un ensemble convexe dans $\mathbb{R}^n$.
L'*intérieur relatif de* X , noté $\operatorname{rint} X$, est défini comme l'intérieur de l'ensemble X lorsque celui-ci est considéré comme une partie de son enveloppe affine $\operatorname{aff} X$ avec la topologie induite.
Plus formellement,

$$\operatorname{rint} X = \{x \in \operatorname{aff} X \ \text{ tel que } \exists\, \varepsilon > 0 \ : (x + \varepsilon \overline{B}) \cap \operatorname{aff} X \subset X\}$$

où $\overline{B}$ est la boule unité fermée de $\mathbb{R}^n$.

1.5 Transformée de Fenchel d'une fonction convexe

Définition 1.31
Soit $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ une fonction convexe.
On appelle *transformée de Fenchel de* f ou *fonction conjuguée de* f la fonction $f^0 : \mathbb{R}^n \to \mathbb{R} \cup \{-\infty, +\infty\}$ définie pour tout $y \in \mathbb{R}^n$ comme

$$f^0(y) = \sup_{x \in \mathbb{R}^n} \{y^T x - f(x)\} .$$

Définition 1.32
L'*inégalité de Fenchel* associée à f^0 est :

$$y^T x \leq f(x) + f^0(y) \qquad \forall\, y, x \in \mathbb{R}^n .$$

1.6 Applications multivoques algorithmiques

Dans les prochains chapitres, nous allons développer divers algorithmes qui comme annoncé dans l'introduction seront basés sur, d'une part, une application multivoque D de construction d'une direction de descente et, d'autre part, sur une application multivoque M de recherche linéaire exacte le long d'une direction de descente relativement à une fonction ψ . Nous allons définir chacune d'entre elles.

1.6.1 Application multivoque $D : x \rightsquigarrow D(x)$

Soient X un ensemble convexe fermé non vide de $\mathbb{R}^n$ et ψ une fonction continue de X dans $\mathbb{R}$.
L'application multivoque qui permet de trouver une direction de descente $d(x)$ pour ψ en x est définie par :

$$\begin{array}{rccc} D: & X & \to & X \times \mathbb{R}^n , \\ & x & \rightsquigarrow & (x, d(x)) . \end{array}$$

Cette application multivoque est souvent définie à partir d'un sous-problème à résoudre qui nous permet de trouver $y(x) \in X$ tel que

$$d(x) = y(x) - x \qquad \forall\, x \in X$$

(cfr. (2.4), ($\mathrm{NNLP}^k_{\varphi^k}$) à la section 2.2.2, ($\mathrm{AVIP}^k_{\Phi^k}$) à la section 3.2.1, ...).

1.6.2 Application multivoque $M : x \rightsquigarrow M(x)$

Soit X un ensemble convexe fermé non vide de $\mathbb{R}^n$.
Lorsque, dans un algorithme, à partir de x , nous voulons déterminer l'itéré suivant x^+ , nous pouvons effectuer une recherche linéaire exacte. Cela consiste à résoudre un problème de minimisation :

$$\min_{t \in L} f(x + td)$$

où L est un sous-ensemble borné de $\mathbb{R}$.
À ce problème de minimisation, nous associons une application multivoque

$$\begin{array}{rccc} M: & X \times \mathbb{R}^n & \to & \mathbb{R}^n \\ & (x, d(x)) & \rightsquigarrow & M(x, d(x)) \end{array}$$

avec

$$M(x, d(x)) = \{x^+ \mid x^+ = x + t^+ d\ ,\ t^+ \in L \text{ et } f(x^+) \leq f(x + td) \quad \forall\, t \in L\}\ .$$

Ces définitions sont tirées de [2].

Chapitre 2

Optimisation non différentiable

2.1 Introduction

Notre but est de proposer un cadre algorithmique pour résoudre un problème d'inéquation variationnelle. Sous sa forme générale, celui-ci s'exprime comme :

(GVIP) $\boxed{\begin{array}{l}\text{Trouver } x^* \in X \text{ et } u^* \in U(x^*) \text{ tels que} \\ \qquad [F(x^*) + u^*]^T \ (x - x^*) \geq 0 \qquad \forall\, x \in X\,,\end{array}}$

où

- X est un ensemble convexe fermé non vide de $I\!R^n$;
- $F : X \to I\!R^n$ est une fonction quelconque;
- $U : X \to 2^{I\!R^n}$ est un opérateur multivoque monotone.

Ce problème est connu sous le nom de *problème d'inéquation variationnelle généralisé*. Ce dernier contient divers cas particuliers. Dans ce mémoire, nous avons choisi d'en développer deux. Le premier d'entre eux porte le nom de problème d'optimisation non linéaire non différentiable sans contrainte. C'est celui-ci que nous allons exposer dans ce chapitre. Nous allons tout d'abord détailler l'algorithme d'approximation du coût qui lui est associé. Nous établirons ensuite la convergence de cet algorithme. Finalement, nous expliquerons le principe du problème auxiliaire et montrerons que le processus itératif qu'il génère est équivalent à l'algorithme sus-mentionné.

Le problème que nous allons étudier se déduit de (GVIP) en choisissant :

- $U = \partial u$ où $u : I\!R^n \to I\!R \cup \{+\infty\}$ est une fonction convexe, propre, semi-continue inférieurement (s.c.i.) et telle que $\mathrm{rint}(\mathrm{dom}\, u) \neq \emptyset$;
- $F = \nabla f$ où $f : I\!R^n \to I\!R \cup \{+\infty\}$ est une fonction de classe C^1 sur $\mathrm{dom}\, u$;
- $X = I\!R^n$.

Nous obtenons ainsi le *problème d'inéquation variationnelle* suivant :

(VIP) $\boxed{\begin{array}{l}\text{Trouver } x^* \in I\!R^n \text{ et } \xi_u^* \in \partial u(x^*) \text{ tels que} \\ \qquad [\nabla f(x^*) + \xi_u^*]^T \ (x - x^*) \geq 0 \qquad \forall\, x \in I\!R^n .\end{array}}$

Le problème peut encore s'exprimer comme :

$$\text{Trouver } x^* \in I\!R^n \ \text{ tel que } \ -\nabla f(x^*) \in \partial u(x^*) . \tag{2.1}$$

Ces formulations sont équivalentes.
En effet, sous l'hypothèse que $\mathrm{rint}(\mathrm{dom}\, u) \cap \mathrm{rint}(X) \neq \emptyset$,

$$\text{(GVIP)} \quad \Leftrightarrow \quad 0 \in F(x^*) + U(x^*) + N_X(x^*) .$$

Dans notre cas particulier,

$$\begin{aligned} \text{(VIP)} \quad &\Leftrightarrow \quad 0 \in \nabla f(x^*) + \partial u(x^*) \\ &\Leftrightarrow \quad -\nabla f(x^*) \in \partial u(x^*) . \end{aligned}$$

■

Nous supposons que les hypothèses sur X , f et u sont valables tout au long du chapitre. Nous admettons également que le problème (VIP) possède toujours au moins une solution. Nous notons Ω l'ensemble de ces solutions.

Exprimons à présent le problème (2.1) sous la forme d'un problème d'optimisation :

(NNLP) $\boxed{\begin{array}{l}\text{Trouver } x^* \in \operatorname*{argmin}_{x \in I\!R^n} T(x) \text{ où} \\ \qquad T(x) \triangleq f(x) + u(x) \qquad \forall\, x \in I\!R^n .\end{array}}$

Ce problème est appelé *programme non différentiable non linéaire sans contrainte.*

Proposition 2.1
(NNLP) *et* (2.1) *sont deux problèmes équivalents.*

Preuve :
Nous remarquons tout d'abord que u est continue sur $\mathrm{rint}(\mathrm{dom}\, u)$ [19, Th. 10.1] et Lipschitz continue sur tout ensemble D , fermé, borné, $D \subseteq \mathrm{rint}(\mathrm{dom}\, u)$ [19, Th. 10.4].

Condition nécessaire pour (NNLP) :

- Lorsque $\mathrm{dom}\, u = I\!R^n$, u est localement lipschitzienne et comme f est continûment différentiable sur $\mathrm{dom}\, u$, il s'en suit qu'elle est localement lipschitzienne. Donc, $T \triangleq f + u$ est localement Lipschitz continue. Par [4], nous affirmons que T possède un gradient généralisé, noté $\partial^* T$ et défini par :
$$\partial^* T(x) = \nabla f(x) + \partial u(x) \qquad \forall\, x \in I\!R^n .$$
Ainsi, par [5], nous obtenons que (2.1) est une condition nécessaire pour (NNLP).

- Lorsque $\mathrm{dom}\, u \neq I\!R^n$, nous devons ajouter une hypothèse pour que le problème (2.1) soit une condition nécessaire pour (NNLP). Ceci est démontré dans [10].

Voici l'hypothèse en question :

Hypothèse 2.1
$u'(x; \cdot)$ est semi-continue inférieurement sur $\mathrm{dom}\, u$.
Par la proposition 1.8, ceci s'écrit sous la forme
$$u'(x;d) = \liminf_{e \to d} u'(x;e) \quad \forall\, x \in \mathrm{dom}\, u \text{ et } \forall\, d \in I\!R^n .$$

Condition suffisante pour (NNLP) :
Si f est convexe, (2.1) est une condition suffisante pour (NNLP). ∎

Remarque 2.1 D'après [18] et [19, Th. 23.2], l'hypothèse 2.1 permet de calculer la dérivée directionnelle de u comme
$$u'(x;d) = \sup\{\xi_u^T d \mid \xi_u \in \partial u(x)\} \tag{2.2}$$
$\forall\, x \in \mathrm{dom}\, u$ et $\forall\, d \in I\!R^n$.

À partir de maintenant, nous supposons que les hypothèses nécessaires à l'équivalence de ces problèmes sont vérifiées. Nous ne distinguons plus (VIP), (2.1) et (NNLP).

2.2 Algorithme d'approximation du coût

Comme expliqué dans l'introduction, l'algorithme d'approximation du coût est un algorithme d'approximations successives basé sur deux grandes idées. À chaque itération, nous devons premièrement déterminer une direction de recherche en résolvant un sous-problème "approximé" du problème initial (VIP). Deuxièmement, nous devons effectuer une recherche linéaire le long de cette direction relativement à une fonction de mérite.

Nous avons déjà défini intuitivement les fonctions de mérite dans l'introduction mais, afin de mieux comprendre la suite du mémoire, formalisons quelque peu cette notion.

Définition 2.1
Une *fonction de mérite* pour le problème (VIP) est une fonction dont les minima coïncident avec les solutions du problème (VIP) .

Une fonction de mérite fréquemment utilisée pour tester la convergence d'un algorithme itératif développé pour résoudre (VIP) est la fonction

$$\psi(x) = \min_{x^* \in \Omega} \|x - x^*\|$$

où $\|\cdot\|$ est la norme euclidienne.
On constate qu'un minimum de la fonction ψ correspond à une de ses racines, c'est-à-dire à $x^* \in \Omega$, une solution de (VIP) .

Cette notion étant précisée, on voit clairement que les relations introduites ci-dessus entre les problèmes (2.1) et (NNLP) nous incitent à choisir T comme fonction de mérite.

Exposons à présent le sous-problème qui nous permettra de générer une direction de recherche. Nous allons approximer ∇f par une fonction monotone d'approximation du coût notée $\Phi : I\!R^n \to I\!R^n$. Comme en $x \in I\!R^n$, la fonction ∇f est remplacée par Φ , nous commettons une erreur d'approximation égale à $\nabla f - \Phi$. Nous la prenons en compte en ajoutant au coût original $\Phi + \partial u$, le terme fixe $\nabla f(x) - \Phi(x)$. Nous obtenons ainsi le sous-problème "approximé" du problème initial :

$$\text{Trouver } y \in X \text{ tel que } 0 \in (\Phi(y) + \partial u(y) + \nabla f(x) - \Phi(x)) \; . \tag{2.3}$$

2.2.1 Algorithme

Voyons précisément comment ce sous-problème va contribuer à résoudre l'inéquation variationnelle initiale en développant l'algorithme d'approximation du coût.

Soit une suite $\{\Phi^k\}$ de fonctions monotones approximant le coût.

Pas 0 (Initialisation)

Choisir $x^0 \in I\!R^n$ et poser $k = 0$.

Pas 1 (Construction d'une direction de recherche à l'aide du sous-problème)

Trouver y^k qui vérifie

$$0 \in \left(\Phi^k(y^k) + \partial u(y^k) + \nabla f(x^k) - \Phi^k(x^k)\right) . \tag{2.4}$$

Poser $d^k = y^k - x^k$.

Pas 2 (Premier critère d'arrêt)

Si x^k satisfait (2.4) , alors STOP : $x^k \in \Omega$.

Sinon, on continue.

Pas 3 (Recherche linéaire le long de la direction d^k relativement à T)

Trouver une longueur de pas t_k qui résout le problème suivant :

$$t_k \in \operatorname*{argmin}_{t \geq 0} T(x^k + td^k) .$$

Pas 4 (Itération)

Poser $x^{k+1} = x^k + t_k d^k$ et $k = k + 1$.

Pas 5 (Second critère d'arrêt)

Si x^k est acceptable comme solution, alors STOP.
Sinon, allez au Pas 1.

Les critères d'arrêt typiques sont une petite longueur de pas OU une petite décroissance dans la valeur de la fonction objectif entre deux itérations.

2.2.2 Reformulation du sous-problème d'approximation

Comme il est plus aisé de résoudre un problème d'optimisation qu'un problème de la forme (2.4), nous allons dériver le problème de minimisation associé à (2.4).

$(\mathrm{NNLP}^k_{\varphi^k})$

Soit $\varphi^k : I\!R^n \to I\!R$ une fonction convexe de classe C^1 sur $\mathrm{dom}\, u$.
Trouver $y^k \in \underset{x \in I\!R^n}{\operatorname{argmin}}\, T_{\varphi^k}(x)$ où

$$\begin{aligned} & T_{\varphi^k}(x) \\ = \quad & \varphi^k(x) + u(x) + f(x^k) - \varphi^k(x^k) + \left[\nabla f(x^k) - \nabla \varphi^k(x^k)\right]^T \, (x - x^k) \, . \end{aligned}$$

L'ensemble, parfois vide, des solutions optimales de $(\mathrm{NNLP}^k_{\varphi^k})$ est noté $Y(x^k)$. Résoudre (VIP) en appliquant l'algorithme d'approximation du coût revient à résoudre une séquence de problèmes d'optimisation $(\mathrm{NNLP}^k_{\varphi^k})$. En effet, $(\mathrm{NNLP}^k_{\varphi^k})$ est équivalent à (2.4) pour tout k .

Preuve :
Soit $x^k \in X$ donné.

$$\begin{aligned} (\mathrm{NNLP}^k_{\varphi^k}) \quad & \Leftrightarrow \quad \text{Trouver } y^k \text{ qui appartient à } \underset{x \in I\!R^n}{\operatorname{argmin}}\, T_{\varphi^k}(x) \\ & \Leftrightarrow \quad \text{Trouver } y^k \text{ tel que } 0 \in \partial T_{\varphi^k}(y^k) \end{aligned}$$

par la proposition 1.17

$$\Leftrightarrow \quad \text{Trouver } y^k \text{ tel que } 0 \in \nabla \varphi^k(y^k) + \partial u(y^k) + \nabla f(x^k) - \nabla \varphi^k(x^k)$$

par [19, p. 223] car les fonctions sont propres et convexes et
$\mathrm{rint}(\mathrm{dom}\, u) \cap \mathrm{rint}(\mathrm{dom}\, \varphi^k) \neq \emptyset$

$$\Leftrightarrow \quad \text{Trouver } y^k \text{ tel que } 0 \in \nabla f(x^k) - \Phi^k(x^k) + \Phi^k(y^k) + \partial u(y^k)$$

car $\Phi^k \equiv \nabla \varphi^k \quad \forall\, k$

$$\Leftrightarrow \quad (2.4)\, .$$

■

2.2.3 Justification de l'algorithme

Théorème 2.1 (Caractérisation du point fixe de Ω)
Soient $\varphi : I\!R^n \to I\!R$ une fonction convexe de classe C^1 sur $\operatorname{dom} u$, $x \in \operatorname{dom} u$, $y \in Y(x)$, $d = y - x$. *Nous supposons également que l'hypothèse 2.1 est vérifiée.*
Alors

$$x \in \Omega \quad \Leftrightarrow \quad x \in Y(x) \quad \Leftrightarrow \quad T_\varphi(x) = T_\varphi(y) .$$

Preuve :

Montrons que $x \in \Omega \Leftrightarrow x \in Y(x)$:

$$\begin{aligned} x \in \Omega \quad & \Leftrightarrow \quad -\nabla f(x) \in \partial u(x) \\ & \Leftrightarrow \quad -\nabla\varphi(x) - \nabla f(x) + \nabla\varphi(x) \in \partial u(x) \\ & \Leftrightarrow \quad x \text{ est solution de (2.4) avec } \Phi = \nabla\varphi \text{ et } \Phi = \Phi^k \quad \forall\, k \\ & \Leftrightarrow \quad x \text{ est solution de } (\mathrm{NNLP}_\varphi) \text{ car } \varphi \text{ est convexe} \\ & \Leftrightarrow \quad x \in Y(x) . \end{aligned}$$

Montrons que $x \in Y(x) \Leftrightarrow T_\varphi(x) = T_\varphi(y)$:

$$\begin{aligned} x \in Y(x) \quad & \Leftrightarrow \quad x \text{ est solution optimale de } (\mathrm{NNLP}_\varphi) \\ & \Leftrightarrow \quad T_\varphi(x) \text{ est la valeur optimale de } T_\varphi . \end{aligned}$$

Or, on sait par hypothèse que $y \in Y(x)$, donc $T_\varphi(y)$ est la valeur optimale de T_φ .
Donc, $x \in Y(x) \Leftrightarrow T_\varphi(x) = T_\varphi(y)$. ∎

Ce théorème justifie le premier critère d'arrêt de l'algorithme. Si $x \in Y(x)$, alors x est solution de (VIP) et nous avons atteint notre but.

Théorème 2.2 (Théorème de descente)
Soient $\varphi : I\!R^n \to I\!R$ une fonction convexe de classe C^1 sur $\operatorname{dom} u$, $x \in \operatorname{dom} u$, $y \in Y(x)$, $d = y - x$. *Nous supposons en outre que l'hypothèse 2.1 est vérifiée.*
Soit $\overline{x} \in I\!R^n$ tel que $T_\varphi(\overline{x}) < T_\varphi(x)$. Alors $T'(x; \overline{x} - x) < 0$.
En particulier pour $\overline{x} = y$, si $x \notin \Omega$, alors $T'(x; d) < 0$.

Preuve :
Soient $x \in \operatorname{dom} u$ et $\overline{x} \in I\!R^n$.

Montrons que $T'(x;\overline{x}-x) < 0$ lorsque $T_\varphi(\overline{x}) < T_\varphi(x)$:
Si on définit $T_\varphi(y)$ comme

$$T_\varphi(y) = \varphi(y) + u(y) + f(x) - \varphi(x) + [\nabla f(x) - \nabla\varphi(x)]^T(y-x) \ ,$$

(a) $T_\varphi(\overline{x}) < T_\varphi(x) \Leftrightarrow \varphi(\overline{x}) - \varphi(x) - \nabla\varphi(x)^T(\overline{x}-x) < u(x) - u(\overline{x}) - \nabla f(x)^T(\overline{x}-x)$.

(b) De plus, φ est convexe par hypothèse, ce qui selon la proposition 1.1 se traduit par :

$$\varphi(\overline{x}) - \varphi(x) \geq \nabla\varphi(x)^T(\overline{x}-x) \ .$$

(c) Ainsi, en rassemblant (a) et(b), nous obtenons :

$$u(x) - u(\overline{x}) > \nabla f(x)^T(\overline{x}-x) \ .$$

De plus,

$$\begin{aligned}
T'(x;\overline{x}-x) &= \lim_{\substack{t\to 0\\ >}} \frac{T(x+t(\overline{x}-x)) - T(x)}{t} \\
&= \lim_{\substack{t\to 0\\ >}} \frac{f(x+t(\overline{x}-x)) - f(x)}{t} + \lim_{\substack{t\to 0\\ >}} \frac{u(x+t(\overline{x}-x)) - f(x)}{t} \\
&= \nabla f(x)^T(\overline{x}-x) + u'(x;\overline{x}-x)
\end{aligned}$$

et par la remarque 2.1 :

$$= \nabla f(x)^T(\overline{x}-x) + \sup\{\xi_u^T(\overline{x}-x) \mid \xi_u \in \partial u(x)\} \ .$$

Par (c),

$$T'(x:\overline{x}-x) < u(x) - u(\overline{x}) + \sup\{\xi_u^T(\overline{x}-x) \mid \xi_u \in \partial u(x)\} \ .$$

Comme u est convexe, $u(\overline{x}) - u(x) \geq \xi_u^T(\overline{x}-x)$ pour tout $\xi_u \in \partial u(x)$, nous avons

$$u(x) - u(\overline{x}) + \xi_u^T(\overline{x}-x) \leq 0 \qquad \forall\, \xi_u \in \partial u(x) \ .$$

Donc,

$$T'(x;\overline{x}-x) < 0 \ .$$

Montrons que, pour $\overline{x} = y$, si $x \notin \Omega$, alors $T'(x;d) < 0$:

(d) $$x \notin \Omega \;\Leftrightarrow\; x \notin Y(x) \Leftrightarrow T_\varphi(x) \neq T_\varphi(y)$$

Comme $y \in Y(x)$, y résout (NNLP$_\varphi$) , i.e. $y \in \operatorname*{argmin}_{x \in I\!R^n} T_\varphi(x)$.

Donc,
$$T_\varphi(y) \leq T_\varphi(x) .$$

Comme, par (d), $T_\varphi(x) \neq T_\varphi(y)$, nous avons
$$T_\varphi(y) < T_\varphi(x) .$$

Donc, $T'(x;d) < 0$ par la première partie de la thèse. ∎

Ce théorème justifie l'algorithme dans ce sens où il affirme que, si $x \notin \Omega$, d est une direction de descente pour T en x ($x \in I\!R^n$). Ainsi, si $x^k \notin \Omega$, lorsque l'algorithme fournit un nouvel itéré x^{k+1} , la valeur de T en ce point est inférieure à celle qu'il possédait en x^k . Nous espérons, grâce à cette propriété, arriver au minimum de T et ainsi à une solution de (NNLP) .

Théorème 2.3 (Existence et unicité des solutions de (NNLP$_\varphi$))
Sous les mêmes hypothèses que le théorème de descente,

1. *si φ ou u est strictement convexe sur* $\operatorname{dom} u$*;*

   si $Y(x)$ est non vide,

   alors $Y(x)$ est un singleton.

2. *si φ est fortement convexe sur* $\operatorname{dom} u$ *,*

   alors $Y(x)$ est non vide et est un singleton et
$$T'(x;d) \leq -m_\varphi \|d\|^2 .$$

Preuve :

1. φ ou u est strictement convexe sur $\operatorname{dom} u$. Donc T_φ est strictement convexe sur $\operatorname{dom} u$ (cfr. annexe A.4).

   Comme $Y(x)$ est non vide, T_φ possède au moins un minimum. Mais comme T_φ est strictement convexe, ce minimum est unique par la caractérisation d'unicité d'un minimum au paragraphe 1.2.7.

2. u est convexe et φ est fortement convexe sur $\operatorname{dom} u$. Donc T_φ est fortement convexe sur $\operatorname{dom} u$ (cfr. annexe A.4). $Y(x)$ est donc un singleton puisque (NNLP$_\varphi$) admet un minimum unique (cfr. paragraphe 1.2.7, proposition 1.19).

Montrons que $T'(x;d) \leq -m_\varphi \|d\|^2$ où $x \in \operatorname{dom} u$, $d = y - x$:

$y \in Y(x)$, i.e. y est solution du problème (2.4) car (NNLP$_\varphi$) et (2.4) sont équivalents. Ceci s'écrit comme

$$0 \in \nabla\varphi(y) - \nabla\varphi(x) + \nabla f(x) + \partial u(y)$$

car $\Phi \equiv \nabla\varphi$

$$\begin{aligned} &\Leftrightarrow \quad -\nabla\varphi(y) - \nabla f(x) + \nabla\varphi(x) \in \partial u(y) \\ &\Leftrightarrow \quad [-\nabla\varphi(y) - \nabla f(x) - \nabla\varphi(x)]^T (y - x) \geq u(y) - u(x) \end{aligned} \tag{2.5}$$

car u est une fonction convexe.
De plus,

$$u(y) - u(x) \geq \xi_u(x)^T (y - x) \qquad \forall\ \xi_u(x) \in \partial u(x)$$

ce qui se réécrit comme

$$u(x) - u(y) \leq -\xi_u(x)^T (y - x) \qquad \forall\ \xi_u(x) \in \partial u(x)\ . \tag{2.6}$$

En additionnant (2.5) et (2.6), nous obtenons :

$$[-\nabla\varphi(y) - \nabla f(x) + \nabla\varphi(x) - \xi_u(x)]^T (y - x) \geq 0 \qquad \forall\ \xi_u(x) \in \partial u(x)\ . \tag{2.7}$$

Par (2.7) et la convexité forte de φ , nous avons :

$$\begin{aligned} T'(x; y - x) &= \nabla f(x)^T (y - x) + \sup_{\xi_u(x) \in \partial u(x)} \{\xi_u(x)^T (y - x)\} \\ &\leq [\nabla\varphi(x) - \nabla\varphi(y)]^T (y - x) \\ &\leq -m_\varphi\ \|y - x\|^2\ . \end{aligned}$$

Comme $d = y - x$, nous obtenons la thèse. ∎

2.3 Convergence de l'algorithme d'approximation du coût

Ce paragraphe a pour but d'étudier la convergence de l'algorithme d'approximations successives détaillé dans la section précédente. Avant d'établir la convergence de cet algorithme, deux lemmes techniques sont nécessaires. Le premier d'entre eux concerne la continuité de l'application multivoque définissant la direction de descente à chaque itération (cfr. paragraphe 1.6). Le second traite de la convergence vers un ensemble.

Lemme 2.1 (Continuité de l'application multivoque $x \rightsquigarrow D(x)$)
Soit $\varphi : I\!R^{2n} \rightarrow I\!R$ une fonction continue sur $\mathrm{dom}\, u \times \mathrm{dom}\, u$ *, de la forme $\varphi(x, y)$, convexe et de classe C^1 sur* $\mathrm{dom}\, u$ *par rapport à x . Soit $S \subseteq \mathrm{dom}\, u$ un ensemble arbitraire, fermé dans $I\!R^n$. Nous supposons que l'hypothèse 2.1 est vérifiée.*

1. *Les applications multivoques $x \rightsquigarrow Y(x)$ et $x \rightsquigarrow D(x)$ sont fermées sur S .*

2. *L'ensemble $Y(x)$ est non vide et borné si et seulement si*

$$-\nabla f(x) + \nabla_x \varphi(x, x) \in \mathrm{int}(\mathrm{dom}(u + \varphi(x, \cdot))^0) \tag{2.8}$$

   où $(u + \varphi(x, \cdot))^0$ est la fonction conjuguée de $u + \varphi(x, \cdot)$.

3. *Si, en plus de l'hypothèse (2.8) vérifiée pour tout $x \in S$, $\varphi(x, \cdot)$ est strictement convexe sur S pour tout $x \in S$, ou u est strictement convexe sur S ,*

   alors l'application multivoque $x \rightsquigarrow D(x)$ est continue sur S .

Preuve :

1. **Montrons que l'application multivoque $x \rightsquigarrow Y(x)$ est fermée sur S :**

   Soient $\{x^k\} \in S$ et $\{y^k\} \in \mathrm{dom}\, u$ tels que

$$\begin{array}{lll} \{x^k\} \rightarrow x \,, \quad \{y^k\} \rightarrow y & \text{et} & y^k \in Y(x^k) \quad \forall\, k \\ y^k \in Y(x^k)\ \forall\, k & \Leftrightarrow & -\nabla_x \varphi(x^k, y^k) - \nabla f(x^k) + \nabla_x \varphi(x^k, x^k) \in \partial u(y^k)\ \forall\, k \,. \end{array}$$

   Par [19, Th.24.4] et la continuité de ∇f et $\nabla_x \varphi$ sur $\mathrm{dom}\, u$ et $(\mathrm{dom}\, u \times \mathrm{dom}\, u)$ respectivement, nous pouvons conclure que

$$-\nabla_x \varphi(x, y) - \nabla f(x) + \nabla_x \varphi(x, x) \in \partial u(y)$$

qui, par la convexité de u et $\varphi(x,\cdot)$, donne $y \in Y(x)$.

Donc, selon les définitions 1.20 et 1.21, nous pouvons affirmer que l'application multivoque qui à x associe $Y(x)$ est fermée.

Montrons que l'application multivoque $x \rightsquigarrow D(x)$ est fermée sur S :

$$\operatorname{graph} Y = \{(x,y) \in I\!R^{2n} \mid y \in Y(x)\}$$

désigne le graphe de l'application multivoque $Y : I\!R^n \to 2^{I\!R^n}$ qui définit les solutions du sous-problème (2.4) .

Par [19, Th. 24.4], cet ensemble est fermé dans $I\!R^{2^n}$.

Le graphe de l'opérateur permettant de trouver la direction de recherche est obtenu du graphe de Y par une transformation affine puisque

$$\forall\ x \in I\!R^n \qquad D(x) = Y(x) - x\ .$$

Il est donc toujours fermé. Donc, par [13, p. 65], nous en déduisons que l'application multivoque $x \rightsquigarrow D(x)$ est fermée sur S .

2. Nous admettons ce résultat qui peut être facilement démontré en rassemblant diverses propriétés regroupées dans [19].

3. **Montrons que l'application multivoque $x \rightsquigarrow D(x)$ est continue sur S :**

   Par le théorème d'existence et d'unicité des solutions de (NNLP$_\varphi$), nous pouvons affirmer que $Y(x)$ est un singleton puisque φ ou u est strictement convexe sur S et, par 2., $Y(x)$ est non vide. Donc, l'application $x \rightsquigarrow D(x)$ est univoque sur S puisque, à un x donné, il correspond une et une seule direction $d = y(x) - x$.

   De plus, par la première partie de la thèse, l'application $x \rightsquigarrow D(x)$ est fermée. En rassemblant tous ces éléments, nous pouvons déduire que

$$\text{si } \{x^k\} \underset{(k\to\infty)}{\longrightarrow} x \qquad \text{alors } \{D(x^k)\} \underset{(k\to\infty)}{\longrightarrow} D(x)$$

   ce qui est une caractérisation de la continuité de l'application $x \rightsquigarrow D(x)$ sur S . ∎

Lemme 2.2 (Convergence vers un ensemble)
Soit $S \subseteq C \subseteq I\!R^n$ où S et C sont deux ensembles non vides et C est compact.
Si $\{x^k\}$ est une suite dans C telle que tout point d'accumulation appartient à S ,
alors $\{x^k\}$ converge vers S dans ce sens où

$$\{\inf_{x \in S} \|x^k - x\|\} \underset{(k \to \infty)}{\longrightarrow} 0 \, .$$

Théorème 2.4 (Convergence de l'algorithme d'approximation du coût)
Soient u une fonction continue sur $\operatorname{dom} u$, $\varphi : I\!R^{2^n} \to I\!R$ une fonction continue sur $\operatorname{dom} u \times \operatorname{dom} u$, de la forme $\varphi(x, y)$, convexe et de classe C^1 sur $\operatorname{dom} u$ par rapport à x .
Supposons que x^0 soit choisi tel que l'ensemble de niveau $\mathcal{L}^T_{I\!R^n}(x^0)$ soit borné et que $(NNLP_\varphi)$ soit bien défini (i.e. que $Y(x) \neq \emptyset$ et borné $\forall\ x \in \operatorname{dom} u$).
Supposons en outre que l'hypothèse 2.1 soit vérifiée.
Alors

- $\{T(x^k)\} \to T(\overline{x})$ *avec* $\overline{x} \in \Omega$ *;*
- *tout point d'accumulation x^∞ de $\{x^k\}$ appartient à Ω ; et*
- $\left\{ \inf_{x \in \Omega \cap \mathcal{L}^T_{I\!R^n}(x_0)} \|x^k - x\| \right\} \underset{(k \to \infty)}{\longrightarrow} 0 \, .$

Preuve :
Nous allons essayer de démontrer la thèse en utilisant le Théorème A de Zangwill (paragraphe 1.3). Dans ce but, nous devons identifier

- X à $I\!R^n$,
- Ω à l'ensemble des solutions de (VIP),
- la fonction de descente à T ,
- l'application multivoque A à l'opérateur $M \circ D$ où D est l'application qui définit la direction de recherche et M est celle qui permet la recherche linéaire exacte (cfr. paragraphe 1.6).

Nous devons à présent vérifier que les hypothèses du théorème A de Zangwill sont valables dans notre cas :

1. **Montrons que la suite $\{x^k\} \in E$, E un ensemble compact :**

   $\mathcal{L}^T_{\mathbb{R}^n}(x_0) = \{x \in \mathbb{R}^n \mid T(x) \leq T(x^0)\}$ est borné par hypothèse.

   Comme la direction construite par l'algorithme est une direction de descente (théorème 2.2), $\{x^k\} \in \mathcal{L}^T_{\mathbb{R}^n}(x_0)$.

   De plus, par [24, p. 134], nous savons que $\mathcal{L}^T_{\mathbb{R}^n}(x_0)$ est fermé.

   Donc, $\{x^k\} \in \mathcal{L}^T_{\mathbb{R}^n}(x_0)$ avec $\mathcal{L}^T_{\mathbb{R}^n}(x_0)$ un compact.

2. **Montrons que $T(x^{k+1}) < T(x^k)$ si $x^k \notin \Omega$:**

   Ceci est immédiat par le théorème de descente de la section 2.2.3.

   L'algorithme se termine quand $x^k \in \Omega$.

3. **Montrons que l'application multivoque A est fermée $\forall\ x \notin \Omega$:**

   - Par le lemme 2.1, nous savons que l'application multivoque D est fermée sur $\mathcal{L}^T_{\mathbb{R}^n}(x_0) \subseteq \operatorname{dom} u$.
   - M est fermée sur $\mathcal{L}^T_{\mathbb{R}^n}(x_0) \times \mathbb{R}^n$ par [2, Th. 8.4.1, p. 282] car T est continue sur $\operatorname{dom} u$, $d \neq 0$ et l'ensemble dans lequel t est choisi est fermé.
   - Comme $\mathcal{L}^T_{\mathbb{R}^n}(x_0)$ est borné, [2, cor. 1, p. 252] permet d'affirmer que $A = M \circ D$ est fermée sur $\mathcal{L}^T_{\mathbb{R}^n}(x_0)$.

Nous pouvons à présent appliquer le théorème A de Zangwill et nous obtenons que :

- tout point d'accumulation x^∞ de la suite $\{x^k\}$ appartient à Ω . Un tel point d'accumulation existe car $\{x^k\}$ est bornée.
- $\{T(x^k)\} \to T(\overline{x})$ avec $\overline{x} \in \Omega$ car cette suite est décroissante et minorée.

Le dernier résultat provient du lemme 2.2 avec $C \equiv \mathcal{L}^T_{\mathbb{R}^n}(x_0)$ qui est non vide et $S \equiv \Omega \cap \mathcal{L}^T_{\mathbb{R}^n}(x_0)$ qui est compact et non vide. Pour que toutes les hypothèses soient satisfaites, il faut encore voir que tous les points d'accumumation de C appartiennent à S , ce qui est vérifié car T est continue sur $\operatorname{dom} u$. ∎

2.4 Principe du problème auxiliaire

Rappelons-nous que, dans l'algorithme que nous venons de développer, nous avons généré une suite $\{x^k\}$ d'éléments de $I\!R^n$ telle que tout point d'accumulation est une solution du problème (VIP). Nous avons procédé en deux étapes :

- la recherche d'une direction de descente;
- une recherche linéaire exacte le long de cette direction.

Pour résoudre le problème (VIP), Cohen a, quant à lui, développé une méthode itérative qui ne fait pas appel à une recherche linéaire (cfr. [6]). Voici le problème auxiliaire de Cohen sur lequel se base son algorithme :

Soit $\varphi^k : I\!R^n \longrightarrow I\!R$ une fonction convexe de classe C^1 sur $\mathrm{dom}\, u$, $\forall\, k$.
Étant donné $x^k \in I\!R^n$, trouver $x^{k+1} \in I\!R^n$ tel que

$$x^{k+1} \in \operatorname*{argmin}_{x \in I\!R^n} \left\{ \varphi^k(x) + \varepsilon_k u(x) + [\varepsilon_k \nabla f(x_k) - \nabla \varphi^k(x^k)]^T x \right\}$$

avec $\varepsilon_k > 0$.

Ce problème est équivalent au sous-problème utilisé dans l'algorithme d'approximation du coût pour déterminer la direction de descente. Le mot "*équivalent*" signifie dans ce cas que nous pouvons, par des changements d'écriture, nous ramener à une même formulation pour les deux problèmes.

Si chaque fonction Φ^k approximant ∇f est le gradient d'une fonction convexe φ^k , alors le sous-problème (2.4) utilisé dans l'algorithme d'approximation du coût est équivalent au problème auxiliaire de Cohen.

Preuve :
On sait déjà que, si $\Phi^k = \nabla \varphi^k$ avec φ^k une fonction convexe pour tout k , alors $(\mathrm{NNLP}^k_{\varphi^k})$ est équivalent au problème (2.4). Il reste à montrer que résoudre $(\mathrm{NNLP}^k_{\varphi^k})$ est équivalent à résoudre le problème auxiliaire de Cohen.

$$\begin{aligned} \min_{x \in I\!R^n} T_{\varphi^k}(x) \quad \Leftrightarrow \quad \min_{x \in I\!R^n} & \left\{ \varphi^k(x) + u(x) + f(x^k) - \varphi^k(x^k) \right. \\ & \left. + [\nabla f(x^k) - \nabla \varphi^k(x^k)]^T (x - x^k) \right\} \end{aligned}$$

$$\Leftrightarrow \quad \min_{x \in I\!R^n} \left\{ \frac{\varphi^k(x)}{\varepsilon^k} + u(x) + f(x^k) - \frac{\varphi^k(x^k)}{\varepsilon^k} \right.$$

$$\left. + \left[\nabla f(x^k) - \frac{\nabla \varphi^k(x^k)}{\varepsilon^k} \right]^T (x - x^k) \right\}$$

car on rebaptise $\frac{\varphi^k}{\varepsilon^k}$ par φ^k

$$\Leftrightarrow \quad \min_{x \in I\!R^n} \left\{ \varphi^k(x) + \varepsilon^k u(x) + [\varepsilon^k \nabla f(x^k) - \nabla \varphi^k(x^k)]^T x \right\}$$

car on multiplie par $\varepsilon^k > 0$ et on ne tient plus compte des constantes. ■

2.5 Conclusions

Dans ce chapitre, nous avons exposé deux processus itératifs pour résoudre la première classe de cas particuliers du problème (GVIP) que nous avons appelés *programmes d'optimisation non différentiable non linéaire sans contrainte*. Ces problèmes sont des restrictions de (GVIP) car la fonction F définissant l'inéquation variationnelle n'est pas quelconque mais il s'agit du gradient d'une fonction f . De plus, le problème ne possède aucune contrainte.

Nous avons principalement développé l'algorithme d'approximation du coût. Nous avons vu que la suite d'itérés qu'il engendre possède un point d'accumulation qui est une solution de (VIP) sous les hypothèses suivantes : u est continue, φ est convexe et continue, $\mathcal{L}^T_{I\!R^n}(x^0)$ est borné, $Y(x)$ est non vide et borné. Finalement, nous avons exposé le second algorithme basé sur le principe du problème auxiliaire. Celui-ci se différencie du premier par le fait qu'on n'y effectue pas de recherche linéaire. Notons cependant que les deux processus itératifs sont équivalents à un changement d'écriture près.

Nous accorderons davantage d'attention au principe du problème auxiliaire dans le chapitre 4 où nous étudierons en détail sa convergence sous l'hypothèse de co-coercivité. Dans les prochains chapitres, nous porterons attention aux méthodes algorithmiques de résolution d'inéquations variationnelles dans le cadre plus général où la fonction F ne sera plus exprimée comme le gradient d'une fonction f .

Chapitre 3

Inéquations variationnelles non symétriques

3.1 Introduction

Dans le chapitre précédent, nous avons étudié le premier cas particulier de (GVIP) où $F = \nabla f$ avec $f : I\!R^n \rightarrow I\!R \cup \{+\infty\}$. À présent, nous allons nous intéresser au second qui est caractérisé par le fait que la fonction F ne peut pas s'écrire comme le gradient d'une fonction f quelle qu'elle soit. Par opposition au problème résolu dans le deuxième chapitre, nous l'appellerons ici *problème d'inéquation variationnelle non symétrique* (AVIP) . Il s'agit d'une particularisation de (GVIP) car

- X est un ensemble convexe fermé non vide de $I\!R^n$,
- $F : X \rightarrow I\!R^n$ est une fonction continue sur X ,
- $U \equiv \partial u$ avec $u : I\!R^n \rightarrow I\!R \cup \{+\infty\}$ une fonction convexe, propre, s.c.i.

Le problème d'inéquation variationnelle non symétrique se définit comme :

(AVIP)

Trouver $x^* \in X$ et $\xi_u^* \in \partial u(x^*)$ tels que
$$[F(x^*) + \xi_u^*]^T (x - x^*) \geq 0 \qquad \forall\, x \in X \, .$$

Dans ce chapitre, nous nous proposons de suivre un schéma analogue à celui emprunté pour résoudre les problèmes non différentiables. Nous allons tout d'abord exposer un algorithme d'approximation du coût avec une longueur de pas fixée car, comme nous le verrons, le recherche linéaire est impossible puisque aucune fonction de mérite n'est connue. Afin de remédier à ce problème, nous construirons une classe de fonctions de mérite pour (AVIP) et nous en étudierons diverses propriétés. Ensuite, nous présenterons un algorithme de descente avec recherche linéaire exacte et en étudierons la convergence. Finalement, nous exposerons le principe du problème auxiliaire associé à (AVIP).

Pour faciliter l'analyse du problème (AVIP) , nous nous proposons de le formuler de manière équivalente en introduisant la fonction u au lieu de son sous-différentiel. Ceci est possible moyennant une hypothèse de régularité que nous supposons vérifiée tout au long du chapitre.

Hypothèse 3.1
$\mathrm{int}(\mathrm{dom}\, u) \cap X \neq \emptyset$.

Rockafellar a montré dans [19] et [23] que cette hypothèse pouvait en général s'écrire sous la forme suivante :

Hypothèse 3.2
$\mathrm{rint}(\mathrm{dom}\, u) \cap \mathrm{rint}\, X \neq \emptyset$.

Cette hypothèse est nécessaire pour que l'identité suivante soit correcte [19, Th. 23.8] :

$$\partial(\delta_X + u)(x) = \partial\delta_X(x) + \partial u(x) \ ,$$

où δ_X est la fonction indicatrice de l'ensemble X définie par

$$\delta_X(x) = \begin{cases} 0 & \text{si } x \in X \ , \\ +\infty & \text{sinon.} \end{cases}$$

Par [18], [20], [21], [19, Th. 27.4] et l'hypothèse (3.1), (AVIP) est équivalent au problème suivant :

$$\boxed{\begin{array}{ll} \text{Trouver } x^* \in X \text{ tel que} & \\ F(x^*)^T(x - x^*) + u(x) - u(x^*) \geq 0 & \forall\, x \in X \ . \end{array}} \tag{3.1}$$

Nous attirons votre attention sur le fait que, comme (AVIP) et (3.1) sont des problèmes équivalents, nous étudierons leurs propriétés sans les distinguer.

Nous supposons que (AVIP) admet toujours au moins une solution. Nous notons Ω l'ensemble des solutions de (AVIP) .
Selon [7, Th. 3.1], une condition suffisante pour que Ω soit toujours non vide est que F soit monotone sur X et que l'une des hypothèses suivantes soit vérifiée :

- $\operatorname{dom} u \cap X$ est borné (en supposant que $\operatorname{int}(\operatorname{dom} u) \cap X \neq \emptyset$);
- $\exists\, \overline{x} \in \operatorname{dom} u \cap X$ tel que

$$\lim_{\substack{x \in X \\ \|x\| \to +\infty}} \left\{ \frac{F(x)^T(x - \overline{x}) + u(x)}{\|x\|} \right\} = +\infty \,.$$

De plus, Ω est un singleton, i.e. (AVIP) admet une unique solution si Ω est non vide et si une des conditions suivantes est vérifiée :

- F est strictement monotone sur X (cfr. définition 1.5);
- u est strictement convexe sur X (cfr. définition 1.2).

Nous supposons également que les hypothèses sur X , F et u restent valables tout au long du chapitre.

Du fait que F ne peut plus s'écrire comme le gradient d'une fonction f , (AVIP) ne peut pas être converti sans ambiguïté en un problème d'optimisation comme ce fut le cas dans le chapitre précédent. Une fonction de mérite utilisable dans un algorithme d'approximation du coût n'est donc pas perceptible. Nous verrons cependant dans la suite du chapitre que le sous-problème que nous allons développer dans l'algorithme d'approximation du coût va induire une classe de fonctions de mérite.

3.2 Algorithme d'approximation du coût avec longueur de pas fixée

Comme nous n'avons pas encore défini de fonction de mérite, nous présentons dans ce paragraphe un algorithme d'approximations successives dans lequel nous ne faisons pas de recherche linéaire pour trouver la longueur de pas. Nous la fixons donc égale à l'unité.

3.2.1 Algorithme d'approximation du coût

Soit une suite $\{\Phi^k\}$ de fonctions monotones approximant le coût.

Pas 0 (Initialisation)

Choisir un point $x^0 \in X$.

Poser $k = 0$.

Pas 1 (Direction de recherche)

(AVIP$^k_{\Phi^k}$)

Trouver $y^k \in X$ et $\xi^k_u \in \partial u(y^k)$ tels que
$[\Phi^k(y^k) + F(x^k) + \xi^k_u - \Phi^k(x^k)]^T(y - y^k) \geq 0 \;\; \forall\, y \in X$.

Poser $d^k = y^k - x^k$.

Pas 2 (Critère d'arrêt)

Si x^k résout (AVIP$^k_{\Phi^k}$) , alors STOP : $x^k \in \Omega$.

Sinon, on continue.

Pas 3 Poser $t_k = 1$.

Pas 4 (Itération)

Poser $x^{k+1} = x^k + t_k d^k = x^k + d^k$.

$k = k + 1$.

Pas 5 (Critère d'arrêt)

Si x^k est acceptable comme solution, alors STOP.

Sinon, aller au pas 1.

3.2.2 Reformulation du sous-problème (AVIP$^k_{\Phi^k}$)

Afin de faciliter la résolution de l'algorithme, nous pouvons introduire une suite $\{\varphi^k\}$ de fonctions convexes telles que $\varphi^k : X \to I\!R$ pour tout k , $\varphi^k \in C^1$ sur X pour tout k et $\nabla\varphi^k = \Phi^k$ pour tout k .
Ainsi, nous réduisons (AVIP$^k_{\Phi^k}$) à un problème d'optimisation appelé *programme non linéaire non différentiable avec contraintes* (CNNLP$^k_{\varphi^k}$) .

(CNNLP$^k_{\varphi^k}$)

Étant donné $x^k \in X$, trouver $y^k \in X$ tel que

$$y^k \in \underset{y\in X}{\operatorname{argmin}}\{\varphi^k(y) - \varphi^k(x^k) + u(y) - u(x^k) \\ +[F(x^k) - \nabla\varphi^k(x^k)]^T(y - x^k)\}$$

Résoudre $(AVIP)$ en appliquant l'algorithme d'approximation du coût revient à résoudre une séquence de sous-problèmes de minimisation (CNNLP$^k_{\varphi^k}$).
En effet, (CNNLP$^k_{\varphi^k}$) est équivalent à (AVIP$^k_{\Phi^k}$) pour tout k .

Preuve :
Soit $x^k \in X$ donné.

(CNNLP$^k_{\varphi^k}$) $\Leftrightarrow$ Trouver y^k tel que

$$y^k \in \underset{y\in I\!R^n}{\operatorname{argmin}}\{\varphi^k(y) - \varphi^k(x^k) + u(y) - u(x^k) \\ +[F(x^k) - \nabla\varphi^k(x^k)]^T(y - x^k) + \delta_X(y)\}$$

$\Leftrightarrow$ Trouver $y^k \in X$ tel que y^k soit un zéro du sous-différentiel de la fonction à minimiser dans l'expression ci-dessus

cfr. proposition 1.17

$\Leftrightarrow$ Trouver $y^k \in X$ tel que

$$0 \in \nabla\varphi^k(y^k) + \partial u(y^k) + F(x^k) - \nabla\varphi^k(x^k) + N_X(y^k)$$

par [19, p. 223] car les fonctions sont convexes, propres et

$\text{rint}(\text{dom}\,\varphi^k) \cap \text{rint}(\text{dom}\,u) \cap \text{rint}(\text{dom}\,\delta_X) \neq \emptyset$ (par l'hypothèse (3.2)

$$\Leftrightarrow \quad \text{Trouver } y^k \in X \text{ et } \xi_u^k \in \partial u(y^k) \text{ tels que}$$

$$\Phi^k(x^k) - \Phi^k(y^k) - F(x^k) - \xi_u^k \in N_X(y^k)$$

car $\Phi^k \equiv \nabla\varphi^k \quad \forall\, k$

$$\Leftrightarrow \quad (\text{AVIP}^k_{\Phi^k}) .$$

■

3.3 Une classe de fonctions de mérite pour (AVIP)

3.3.1 Fonctions de gap

Comme nous l'avons signalé précédemment, F ne peut s'écrire comme le gradient d'aucune fonction f . De ce fait, nous ne pouvons exprimer (AVIP) sous la forme d'un problème de minimisation et nous n'avons pas de choix évident pour la fonction de mérite à utiliser dans l'algorithme. Nous allons donc essayer de construire des fonctions de mérite en nous basant sur la définition suivante :

Définition 3.1
Une fonction $\psi : X \to I\!R \cup \{-\infty, +\infty\}$ est une *fonction de gap pour* (AVIP) si

1. ψ est de signe constant sur X ;

2. $\psi(x) = 0 \Leftrightarrow x \in \Omega$.

De la définition, nous pouvons déduire que ψ donne une mesure de la violation de (AVIP) en tout point de X . Si nous admettons que $\psi(x) \geq 0$, pour tout $x \in X$, en minimisant ψ sur X , nous obtenons un point dans Ω . De ce fait, nous pouvons utiliser les fonctions de gap comme fonctions de mérite pour les inéquations variationnelles. Un exemple de fonction de mérite a été proposé au paragraphe 2.2.

Comment générer une fonction de mérite ?

Nous allons construire une fonction ψ à partir de l'expression à minimiser dans le problème (CNNLP_φ) . Ensuite, nous vérifierons qu'il s'agit bien d'une fonction de gap. Soit $\varphi : I\!R^n \to I\!R \cup \{+\infty\}$ convexe, s.c.i. et $\varphi \in C^1$ sur X .

Soit $L \;:\; X \times X \to I\!R$,

$$(x,y) \leadsto L(x,y) = u(x) - u(y) + \varphi(x) - \varphi(y) + [F(x) - \nabla\varphi(x)]^T(x-y) .$$

Nous définissons $\psi : X \to I\!R^+ \cup \{+\infty\}$ comme

$$\boxed{\psi(x) = \sup_{y \in X} L(x,y)} \tag{3.2}$$

L'ensemble des solutions de (3.2) se note $Y(x)$. Il peut être vide. Il existe différentes formulations équivalentes du problème (3.2) :
Soit $x \in X$ fixé. Trouver $y(x)$ tel que

$$[\nabla\varphi(y(x)) + F(x) - \nabla\varphi(x)]^T(y - y(x)) + u(y) - u(y(x)) \geq 0 \qquad \forall\, y \in X\ , \tag{3.3}$$

et trouver $y(x)$ tel que

$$-\,(F(x) - \nabla\varphi(x)) \in \nabla\varphi(y(x)) + \partial u(y(x)) + N_X(y(x))\ . \tag{3.4}$$

Ces équivalences sont démontrées en annexe (cf. Annexe A.5).

Théorème 3.1
Soit $\varphi : I\!R^n \to I\!R \cup \{+\infty\}$ une fonction convexe, s.c.i. et $\varphi \in C^1$ sur X .
Alors $\psi : X \to I\!R^+ \cup \{+\infty\}$ est une fonction de gap pour (AVIP).

Preuve :
Afin de démontrer la thèse, nous allons utiliser l'inégalité de Fenchel (paragraphe 1.5) définie par

$$h(x) + h^0(z) - z^T x \geq 0 \qquad \forall\, x \in I\!R^n \text{ et } \forall\, z \in I\!R^n \tag{3.5}$$

avec $h : I\!R^n \to I\!R \cup \{+\infty\}$, propre, convexe et h^0 est la fonction conjuguée de h , définie par

$$h^0(w) = \sup_y \{w^T y - h(y)\}\ . \tag{3.6}$$

Pour prouver que $\psi(x) = 0 \Leftrightarrow x \in \Omega$, nous allons choisir

$$\left\{\begin{array}{lcl} h &=& u + \delta_X + \varphi\ , \\ z &=& \nabla\varphi(x) - F(x)\ , \end{array}\right. \tag{3.7}$$

avec $\varphi : I\!R^n \to I\!R \cup \{+\infty\}$, $\delta_X : I\!R^n \to I\!R \cup \{+\infty\}$, $u : I\!R^n \to I\!R \cup \{+\infty\}$.
Ce choix de h convient car :

- $h : I\!R^n \to I\!R \cup \{+\infty\}$;

- h est propre vu que $\mathrm{int}(\mathrm{dom}\, u) \cap X \neq \emptyset$ et pour tout $x \in I\!R^n \quad h(x) > -\infty$ (cfr. définitions 1.16 et 1.17);
- h est convexe car h est la somme de trois fonctions convexes.

1. **Montrons que $\psi(x) \geq 0$ pour tout $x \in X$:**

   Soit $x \in X$ choisi arbitrairement.

$$\begin{aligned}
\psi(x) &= \sup_{y \in X} L(x,y) \\
&= \sup_{y \in X} \{u(x) - u(y) + \varphi(x) - \varphi(y) + [F(x) - \nabla\varphi(x)]^T (x-y)\} \\
&= u(x) + \varphi(x) + \sup_{y \in X} \{-u(y) - \varphi(y) - \delta_X(y) + [F(x) - \nabla\varphi(x)]^T (x-y)\} \\
&= u(x) + \varphi(x) + \delta_X(x) - [\nabla\varphi(x) - F(x)]^T x \\
&\qquad + \sup_{y \in X} \{-u(y) - \varphi(y) - \delta_X(y) + [\nabla\varphi(x) - F(x)]^T y\}
\end{aligned}$$

par (3.7) :

$$= h(x) - z^T x + \sup_{y \in X} \{z^T y - h(y)\}$$

par (3.6) :

$$= h(x) - z^T x + h^0(z) \tag{3.8}$$

   Donc, $\psi(x) \geq 0$ pour tout $x \in X$ par définition de l'inégalité de Fenchel.

2. **Montrons que $\psi(x) = 0 \Leftrightarrow x \in \Omega$:**

   Soit $x \in X$ choisi arbitrairement.

$$\psi(x) = 0 \quad \Leftrightarrow \quad \begin{cases} h(x) + h^0(z) - z^T x = 0 \;, \\ z = \nabla\varphi(x) - F(x) \;, \\ h(x) = (u + \delta_X + \varphi)(x) \;, \end{cases}$$

par [19, Th.23.5, p. 128] :

$$\Leftrightarrow \quad z \in \partial h(x)$$

$$\Leftrightarrow \quad z \in \partial(u + \delta_X + \varphi)(x)$$

par l'hypothèse 3.2 :

$$\Leftrightarrow \quad z \in \partial u(x) + \partial \delta_X(x) + \partial \varphi(x)$$

comme $N_X(x) = \partial \delta_X(x)$:

$$\Leftrightarrow \quad z \in N_X(x) + \partial u(x) + \partial \varphi(x)$$

comme $\varphi \in C^1$ sur X :

$$\Leftrightarrow \quad z \in N_X(x) + \partial u(x) + \nabla \varphi(x)$$

par (3.7) :

$$\Leftrightarrow \quad \nabla \varphi(x) - F(x) \in N_X(x) + \partial u(x) + \nabla \varphi(x)$$

$$\Leftrightarrow \quad -F(x) \in N_X(x) + \partial u(x)$$

par l'annexe A.1 avec $U \equiv \partial u$

$$\Leftrightarrow \quad x \in \Omega\,.$$

■

Nous avons ainsi défini une fonction de mérite pour (AVIP) qui est égale au saut dans l'inégalité de Fenchel pour une fonction h convexe et propre bien choisie.

3.3.2 Fonctions de mérite

Dans ce paragraphe, nous nous proposons d'exposer le rôle des fonctions de mérite ainsi que quelques-unes de leurs propriétés.

L'introduction d'une fonction de mérite possède deux avantages. Elle permet tout d'abord de réaliser une recherche linéaire et donc de construire un algorithme fournissant la (ou les) solution(s) de (AVIP) sans nécessairement fixer la longueur du pas. Elle offre ensuite la possibilité d'établir l'existence et l'unicité des solutions de (AVIP). En effet, le problème de minimisation associé à ψ est

$$(\mathrm{CNNLP}_\psi) \qquad \boxed{\inf_{x \in X} \psi(x)}$$

Comme un minimum de ψ existe toujours si ψ est finie, (CNNLP$_\psi$) peut être utilisé pour détecter la non-existence d'une solution de (AVIP). Si $\min_{x \in X} \psi(x) > 0$, alors $\psi(x) > 0 \quad \forall\, x \in X$. Donc, $\forall\, x \in X \quad x \notin \Omega$. Le problème (AVIP) ne possède pas de solution.

Maintenant que nous possédons une fonction de mérite, nous pouvons l'utiliser dans un algorithme et rechercher l'itété x^{k+1} en faisant une recherche linéaire exacte le long de la direction d^k relativement à ψ . Afin d'assurer la convergence de cet algorithme, nous devons cependant vérifier (par [26, Cor. 4.2.1]) que l'application de recherche linéaire exacte (rappelée au paragraphe 1.6) est fermée. Ceci est possible si la fonction ψ est continue (par [26, Le. 5.1]) sur X . Dans le but de garantir la continuité de ψ , nous supposons à partir de maintenant que la fonction $u : I\!R^n \to I\!R \cup \{+\infty\}$ est continue sur $X \subseteq I\!R^n$. Sous cette hypothèse supplémentaire, Auchmuty a démontré le théorème suivant dans [1] :

Théorème 3.2 (Continuité de ψ)
Si X est borné ou si u ou φ est fortement convexe sur X ,
alors ψ est continue sur X .

Nous fournissons à présent une règle de calcul de la dérivée directionnelle de ψ dans une direction quelconque $d \in I\!R^n$.

Théorème 3.3 (Dérivée directionnelle de ψ)
Soit u une fonction convexe finie sur $I\!R^n$. Soit $F \in C^1$ sur X et $\varphi \in C^2$ sur X .
Sous l'une des deux hypothèses suivantes, on a

$$\begin{aligned} &\forall\, x \in X \quad \text{et} \quad d \in I\!R^n , \\ &\psi'(x;d) = u'(x;d) + [F(x) + (\nabla F(x)^T - \nabla^2\varphi(x))(x - y(x))]^T\, d \end{aligned} \tag{3.9}$$

où $y(x)$ est la solution unique de (3.2).

1. *X est borné et u ou φ est strictement convexe sur X .*

2. *u ou φ est fortement convexe sur X .*

Aux deux propriétés précédentes, nous pouvons ajouter que ψ est une fonction faiblement coercive et qu'elle possède des ensembles de niveau bornés.

Théorème 3.4 (Coercivité faible)
Soient F une fonction fortement monotone de constante m_F sur X et $\varphi \in C^2$ sur X avec un gradient Lipschitz continu de constante $m_{\nabla\varphi}$ sur X .
Si la fonction φ est choisie telle que $M_{\nabla\varphi} < 2m_F$,
alors, $\forall\ x \in X$

$$\|x - x^*\|^2 \leq \frac{2\psi(x)}{2m_F - M_{\nabla\varphi}} \tag{3.10}$$

où x^* est la solution de (3.1).
De plus, la fonction ψ possède des ensembles de niveau bornés et

$$\lim_{\substack{x\in X \\ \|x\|\to+\infty}} \psi(x) = +\infty\ .$$

Preuve :
Soient $x \in X$ quelconque et x^* la solution de (3.1).

1. **Montrons tout d'abord que**

$$\|x - x^*\|^2 \leq \frac{2\psi(x)}{2m_F - M_{\nabla\varphi}}\ .$$

$$\begin{aligned} & F(x)^T(x - x^*) + u(x) - u(x^*) \\ = \ & [F(x)^T(x - x^*) - F(x^*)^T(x - x^*)] + [F(x^*)^T(x - x^*) + u(x) - u(x^*)] \quad \forall\ x \in X\ . \end{aligned}$$

Nous pouvons minorer le premier membre par

$$m_F\|x - x^*\|^2$$

car F est fortement monotone et le second par zéro car x^* est solution du problème (3.1). Ainsi,

$$\forall\ x \in X \qquad F(x)^T(x - x^*) + u(x) - u(x^*) \geq m_F\|x - x^*\|^2\ . \tag{3.11}$$

nous savons que

$$\psi(x) = u(x) - u(y(x)) + \varphi(x) - \varphi(y(x)) + [F(x) - \nabla\varphi(x)]^T(x - y(x))$$

lorsque $y(x) \in Y(x)$.

Comme $Y(x)$ est l'ensemble des solutions de (3.2),

$$\psi(x) \geq u(x) - u(x^*) + \varphi(x) - \varphi(x^*) + [F(x) - \nabla\varphi(x)]^T(x - x^*)$$

par (3.11) :

$$\geq \varphi(x) - \varphi(x^*) - \nabla\varphi(x)^T(x - x^*) + m_F\|x - x^*\|^2$$

par la proposition 1.7 :

$$\geq -\frac{1}{2}M_{\nabla\varphi}\|x - x^*\|^2 + m_F\|x - x^*\|^2$$

Donc

$$2\psi(x) \geq (2m_F - M_{\nabla\varphi})\|x - x^*\|^2$$

et nous obtenons la thèse par hypothèse puisque $2m_F - M_{\nabla\varphi} > 0$.

2. **Montrons que ψ est faiblement coercive** :

   i.e. par la définition 1.18 que

$$\lim_{\substack{\|x\| \to +\infty \\ x \in X}} \psi(x) = +\infty .$$

   Par (3.10), si $\|x\| \to +\infty$, alors $\psi(x) \to +\infty$.

3. **Montrons que ψ possède des ensembles de niveau bornés** :

   Soit

$$\mathcal{L}_X^{\psi}(x) = \{x \in X \mid \psi(x) \leq c\}$$

   avec $c \in I\!R$ quelconque, un ensemble de niveau de ψ dans X .

   Supposons que $\mathcal{L}_X^{\psi}(x)$ ne soit pas borné, i.e.

$$\forall\ M > 0 \qquad \exists\ \hat{x} \in \mathcal{L}_X^{\psi}(x) \ \text{ tel que } \ \|\hat{x}\| > M .$$

   Alors $\hat{x}$ est tel que $\|\hat{x}\| \to +\infty$.

   Nous aboutissons à une contradiction car ψ est faiblement coercive. ∎

Remarque 3.1
Bien que ψ ne soit, en général, pas convexe, elle possède deux propriétés des fonctions fortement convexes :

- ses ensembles de niveau sont bornés;
- elle est faiblement coercive.

3.4 Algorithme de descente pour (AVIP)

Afin de résoudre (AVIP), nous proposons à présent un algorithme basé sur la recherche de directions de descente relativement à la fonction de mérite ψ définie précédemment.

3.4.1 Algorithme de descente

Soit $\varphi : I\!R^n \to I\!R \cup \{+\infty\}$ une fonction convexe, s.c.i. et de classe C^1 sur X .

Pas 0 (Initialisation)

Choisir $x^0 \in X$ et poser $k = 0$.

Pas 1 (Recherche des directions de descente)

Étant donné x^k , trouver y^k qui donne la solution au problème (3.2) :

$$\sup_{y \in X} L(x^k, y)$$

où

$$L(x^k, y) = u(x^k) + \varphi(x^k) - u(y) - \varphi(y) + [F(x^k) - \nabla\varphi(x^k)]^T(x^k - y) .$$

Poser $d^k = y^k - x^k$.

Pas 2 (Critère d'arrêt)

Si x^k résout (3.2), alors STOP : $x^k \in \Omega$.

Sinon, on continue.

Pas 3 (Recherche linéaire le long de la direction d^k relativement à ψ)

Trouver t_k , la longueur de pas qui résout le problème unidimensionnel

$$t_k \in \operatorname*{argmin}_{0 \leq t \leq 1} \psi(x^k + td^k)$$

où $\psi(x) = \sup_{y \in X} L(x, y) \qquad \forall\, x \in X$.

Pas 4 (Itération)

Poser $x^k = x^k + t_k d^k$ et $k = k + 1$.

Pas 5 (Critère d'arrêt)

Si x^k est acceptable comme solution, alors STOP.

Sinon, aller au pas 1.

3.4.2 Justification de l'algorithme

Théorème 3.5 (Caractérisation du point fixe de Ω)
$\forall\ x \in X \quad Y(x)$ est l'ensemble des solutions de (3.2).
Alors $x \in \Omega \ \Leftrightarrow\ x \in Y(x)$.

Preuve :
On sait par la preuve du théorème 3.1 que

$$\psi(x) = 0 \ \Leftrightarrow\ \nabla\varphi(x) - F(x) \in \partial u(x) - N_X(x) + \nabla\varphi(x)\ . \tag{3.12}$$

De plus, on sait par le paragraphe 3.3.1 que le problème (3.2) est équivalent au problème (3.4) exprimé comme

$$\begin{aligned}&\text{Trouver } y(x) \text{ tel que}\\ &-[F(x) - \nabla\varphi(x)] \in N_X(y(x)) + \partial u(y(x)) + \nabla\varphi(y(x))\ .\end{aligned} \tag{3.13}$$

En rassemblant (3.12) et (3.13), nous pouvons déduire que x résout (3.2) si et seulement si $\psi(x) = 0$, c'est-à-dire $x \in Y(x) \ \Leftrightarrow\ \psi(x) = 0$.
De plus, $\psi(x) = 0 \ \Leftrightarrow\ x \in \Omega$ car ψ est une fonction de gap.
Donc, $x \in \Omega \ \Leftrightarrow\ x \in Y(x)$. ∎

Cette caractérisation justifie le premier critère d'arrêt de l'algorithme de descente. Dès que x est solution du sous-problème (3.2), alors il résout le problème initial d'inéquation variationnelle.

Théorème 3.6 (Existence et unicité des solutions de (3.2))

1. *Si X est borné, alors $\forall\ x \in X \quad Y(x)$ est non vide et borné.*
2. *Sous l'une des hypothèses suivantes, nous obtenons que $\forall\ x \in X \quad Y(x)$ est un singleton :*

1. Comme φ est quadratique,

$$(\nabla\varphi(y(x^*)) - \nabla\varphi(x^*))^T(x^* - y(x^*)) = (x^* - y(x^*))^T\nabla^2\varphi(x^*)(y(x^*) - x^*) .$$

De ce fait, (3.16) se réduit à

$$\begin{aligned} 0 \;\leq\; & u'(x^*; y(x^*) - x^*) + u(x^*) - u(y(x^*)) \\ & -(y(x^*) - x^*)^T\nabla F(x^*)(y(x^*) - x^*) . \end{aligned} \tag{3.17}$$

Comme u est strictement convexe, pour $y(x^*) \neq x^*$,

$$u'(x^*; y(x^*) - x^*) + u(x^*) - u(y(x^*)) < 0 .$$

Comme $\nabla F(x^*)$ est s.d.p.,

$$(y(x^*) - x^*)^T\nabla F(x^*)(y(x^*) - x^*) \geq 0 .$$

Donc,

$$u'(x^*; y(x^*) - x^*) + u(x^*) - u(y(x^*)) - (y(x^*) - x^*)^T\nabla F(x^*)(y(x^*) - x^*) < 0 .$$

Mais, par (3.17), on sait que cette expression doit être positive ou nulle. Pour ne pas avoir de contradiction, on doit nécessairement avoir $y(x^*) = x^*$.
Remplaçons $y(x^*)$ par x^* dans (3.15); nous obtenons

$$F(x^*)^T(y - x^*) + u(y) - u(x^*) \geq 0 \qquad \forall\, y \in X .$$

Donc, x^* résout (3.1).

2. Comme $\nabla F(x^*) - \nabla^2\varphi(x^*)$ est s.d.p.,

$$(x^* - y(x^*))^T[\nabla F(x^*) - \nabla^2\varphi(x^*)](y(x^*) - x^*) \leq 0 .$$

Comme u ou φ est strictement convexe, supposons que ce soit u (on peut faire un raisonnement analogue avec φ),

$$u'(x^*; y(x^*) - x^*) + u(x^*) - u(y(x^*)) < 0$$

tant que $x^* \neq y(x^*)$.
Dans ce cas, φ est convexe, i.e., par la proposition 1.1,

$$[\nabla\varphi(y(x^*)) - \nabla\varphi(x^*)]^T(x^* - y(x^*)) \leq 0 .$$

Donc, (3.16) devient strictement négative. Or, on savait qu'elle était positive ou nulle. Donc, on doit prendre $x^* = y(x^*)$ et on conclut que x^* résout (3.1).

3. Nous partons de l'expression (3.16) et nous essayons de la majorer.

   Comme F est fortement monotone (ou u est fortement convexe), φ est fortement convexe sur X et $\nabla\varphi$ est Lipschitz continu sur X ,

$$\left(\frac{1}{2}m_u + m_F - M_{\nabla\varphi} + m_\varphi\right) \|y(x^*) - x^*\|^2 \leq 0 . \tag{3.18}$$

   Mais, comme $M_{\nabla\varphi} - m_\varphi < m_F + \frac{1}{2}m_u$ par hypothèse, il faut nécessairement que $y(x^*) = x^*$ pour que (3.18) soit vérifié.

   Donc, comme pour les autres classes d'hypothèses, x^* résout (3.1). ∎

Remarque 3.2

Dans l'ensemble 3 d'hypothèses, nous avons soit F fortement monotone de constante m_F , soit u fortement convexe de constante m_u . Mais nous prenons comme convention que, si $F(u)$ est monotone (convexe), alors $m_F = 0$ ($m_u = 0$). Ainsi, nous avons toujours au moins une des deux constantes qui doit être nulle dans (3.18) .

Une conséquence directe de ce théorème est que, si un point admissible x n'est pas solution de (3.1), (c'est-à-dire $x \in X \setminus \Omega$), alors la direction $d = y(x) - x$ est une direction admissible de descente pour ψ . Ce théorème justifie donc le Pas 1 de l'algorithme de descente. Ceci étant un résultat important pour la justification de l'algorithme, nous le réécrivons sous forme d'un théorème.

Théorème 3.8 (Théorème de descente)

Soit u une fonction convexe et finie sur $I\!R^n$. Soient $F \in C^1$ sur X et $\varphi \in C^2$ sur X . Supposons que $x \in X \setminus \Omega$.

Sous l'une des deux classes d'hypothèses suivantes, nous avons :

$$\psi'(x; y(x) - x) < 0 \qquad \forall\, y(x) \in Y(x) .$$

1. *Soient X un ensemble borné, u une fonction strictement convexe sur X , $\nabla F(x)$ s.d.p. et φ une fonction quadratique sur X .*

2. *Soient X un ensemble borné, u ou φ une fonction strictement convexe sur X et $\nabla F(x) - \nabla^2\varphi(x)$ s.d.p.*

Soient F une fonction fortement monotone sur X ou u une fonction fortement convexe sur X , φ une fonction fortement convexe sur X et $\nabla\varphi$ Lipschitz continu sur X . Alors

$$\psi'(x; y(x) - x) \leq -\left(\frac{1}{2}m_u + m_F + m_\varphi - M_{\nabla\varphi}\right) \|y(x) - x\|^2 .$$

3.5 Convergence de l'algorithme de descente

Nous allons à présent montrer que la suite $\{x^k\}$ déterminée par l'algorithme de descente va converger vers la solution recherchée de (AVIP). Mais, avant cela, nous avons besoin du lemme suivant qui va assurer la fermeture des applications multivoques $x \rightsquigarrow Y(x)$ et $x \rightsquigarrow D(x)$. En effet, ces résultats seront utilisés dans la preuve du théorème de convergence de l'algorithme.

Lemme 3.1 (Fermeture)

1. $\forall\ x \in X$ *l'ensemble* $Y(x)$ *est fermé et convexe.*

2. $\forall\ x \in X$ *l'application multivoque* $x \rightsquigarrow Y(x)$ *est fermée.*

3. $\forall\ x \in X$ *l'application multivoque* $x \rightsquigarrow D(x)$ *est fermée.*

Preuve :

1. **Montrons que** $Y(x)$ **est fermé et convexe** :

   La fermeture de $Y(x)$ provient de la continuité de la fonction L (cfr. [8]).

   La convexité de $Y(x)$ provient de la convexité du problème (3.2) (cfr. [11]).

2. **Montrons que l'application multivoque** $x \rightsquigarrow Y(x)$ **est fermée sur** X :

   Soient $\{x^k\}$, $\{y^k\}$ deux suites de X telles que $\{x^k\} \longrightarrow x^\infty$ et $\{y^k\} \longrightarrow y^\infty$, et $y^k \in Y(x^k)\ \forall\ k$.

   $$y^k \in Y(x^k)\ \ \forall\ k \quad \Leftrightarrow \quad y^k \text{ résout (3.2) avec } x^k \text{ donné}$$

   par (3.4) :
   $$\Leftrightarrow \quad -(F(x^k) - \nabla\varphi(x^k)) \in \nabla\varphi(y^k) + \partial u(y^k) + N_X(y^k)\ .$$

   Lorsque $k \longrightarrow +\infty$, par continuité de F et $\nabla\varphi$ sur X , la fermeture de ∂u [19, Th. 24.4] et le fait que $N_X(x) = \partial\delta_X$ [19, p. 215], nous obtenons

   $$-(F(x^\infty) - \nabla\varphi(x^\infty)) \in \nabla\varphi(y^\infty) + \partial u(y^\infty) + N_X(y^\infty)$$

   ce qui est équivalent à $y^\infty \in Y(x^\infty)$.

   D'où la fermeture de $x \rightsquigarrow Y(x)$.

3. **Montrons que l'application multivoque $x \rightsquigarrow D(x)$ est fermée sur X :**

   Le graphe de l'application Y est désigné par

   $$\text{graph}Y = \{(x, y) \in X \times X \mid y \in Y(x)\} .$$

   Par [19, Th. 24.4], cet ensemble est fermé dans $X \times X \subseteq I\!R^{2n}$.

   Le graphe de l'application D est obtenu à partir du graphe de Y par une transformation affine car

   $$D(x) = Y(x) - x \qquad \forall\, x \in X .$$

   Il est donc fermé. Ainsi, par [13, p. 65], nous en déduisons que l'application multivoque D est fermée sur X . ■

Théorème 3.9 (Convergence de l'algorithme de descente)
Soit u une fonction convexe, finie sur $I\!R^n$. Soient $F \in C^1$ sur X et $\varphi \in C^2$ sur X .
Soient x^0 choisi arbitrairement dans X et x^{k+1} déterminé par l'algorithme de descente à partir de x^k .
Sous l'une des classes d'hypothèses suivantes, nous obtenons que (AVIP) possède une solution unique et $\{x^k\}$ converge vers celle-ci.

1. *X est borné, u est strictement convexe sur X , ∇F est s.d.p. sur X et φ est quadratique sur X .*

2. *u est fortement convexe sur X ou F est fortement monotone sur X , φ est fortement convexe sur X et $\nabla\varphi$ est Lipschitz continu sur X . De plus, $M_{\nabla\varphi} - m_\varphi < m_F - \frac{1}{2}m_u$ et soit X est borné, soit F est fortement monotone et $M_{\nabla\varphi} < 2m_F$.*

3. *X est borné, u ou φ est strictement convexe sur X et $\nabla F - \nabla^2\varphi$ est s.d.p. sur X .*

Preuve :
L'idée de la preuve est d'appliquer le théorème A de Zangwill (rappelé au paragraphe 1.3). Dans cette optique, nous devons vérifier que ses hypothèses sont remplies.

1. **La suite $\{x^k\}$ reste dans un ensemble compact :**

   - **Pour les cas 1 et 3 :**

     Nous travaillons en dimension finie et X est fermé et borné par hypothèse. De plus, $x^k \in X \ \forall\ k$.

     D'où $\{x^k\}$ reste dans un ensemble compact.

   - **Pour le cas 2 :**

     Dans le cas où X est borné, nous utilisons les mêmes arguments que pour les cas 1 et 3 .

     Si, par contre, F est fortement monotone et $M_{\nabla\varphi} < 2m_F$, alors on peut conclure que ψ a des ensembles de niveau compacts par le théorème 3.4 dont les hypothèses sont vérifiées.

     Par la propriété de descente de l'algorithme, $x^k \in \mathcal{L}_X^{\psi}(x^0)$ pour tout $k \geq 0$.

     Donc $\{x^k\}$ reste dans un compact.

2. - **Si $x^k \notin \Omega$, alors $\psi(x^{k+1}) < \psi(x^k)$:**

     Suite à la recherche linéaire, x^{k+1} est défini comme

     $$x^{k+1} = x^k + t_k d^k$$

     avec d_k une direction de descente en x^k pour ψ (théorème de descente), et

     $$t_k \in \operatorname*{argmin}_{0 \leq t \leq 1} \psi(x^k + td^k) \ .$$

     Donc, $\psi(x^k + t_k d^k) \leq \psi(x^k + td^k) < \psi(x^k)$.

   - Si $x^k \in \Omega$, alors l'algorithme se termine.

3. **L'application multivoque $A = M \circ D$ qui définit l'algorithme est fermée :**

   - L'application multivoque $x \rightsquigarrow D(x) \equiv Y(x) - x$ où $Y(x)$ est l'ensemble des solutions de (3.2) est fermée par le lemme de fermeture.

   - L'application multivoque M de recherche linéaire exacte (cfr. paragraphe 1.6) est fermée par [26, Le. 5.1, p. 105] car ψ est continue sur X (théorème 3.2) et l'intervalle auquel t doit appartenir est fermé et borné.

Donc, $M_\circ D$ est fermée sur $X \cap \mathcal{L}_X^{\psi}(x^0)$ avec $X \cap \mathcal{L}_X^{\psi}(x^0)$ un ensemble compact (car un fermé dans un compact est un compact) par [26, Cor. 4.2.1, p. 96]. La dernière hypothèse du théorème A de Zangwill est vérifiée.

En appliquant le théorème A de Zangwill, nous pouvons conclure que : tout point d'accumulation de $\{x^k\}$ est solution optimale de (AVIP).
Il existe un tel point d'accumulation car $\{x^k\}$ est borné et par le théorème de Bolzano-Weierstrass, de toute suite bornée, on peut extraire une sous-suite convergente. De plus, il est unique car Ω est un singleton. En effet, les conditions d'existence et d'unicité des solutions de (AVIP) introduites au paragraphe 3.1 sont vérifiées sous chacune des classes d'hypothèses.
Ainsi, $\{x^k\}$ converge vers l'unique solution de (AVIP). ∎

Ainsi, nous venons de montrer que, partant de n'importe quel point admissible et utilisant une recherche linéaire exacte, la suite $\{x^k\}$ des itérés définit une suite $\{\psi(x^k)\}$ décroissante et converge vers la solution unique x^* de (AVIP).

3.6 Principe du problème auxiliaire

Comme nous l'avons expliqué dans le chapitre 2, Cohen va résoudre le problème d'inéquation variationnelle en utilisant un processus itératif qui ne nécessite par de recherche linéaire (cfr. [6]). Chaque itération de cet algorithme se base sur le problème auxiliaire suivant :

Soient $\varphi^k : X \longrightarrow I\!R$ une fonction fortement convexe de classe C^1 sur X et $\varepsilon^k > 0$ pour tout k .
Étant donné $x^k \in X$, trouver x^{k+1} tel que

$$x^{k+1} \in \min_{x \in X} \left\{ \varphi^k(x) + [\varepsilon_k F(x^k) - \nabla \varphi^k(x^k)]^T x + \varepsilon_k u(x) \right\} .$$

Le problème de Cohen est équivalent au sous-problème utilisé dans l'algorithme d'approximation du coût employé pour résoudre (AVIP). En effet,

$$\min_{x \in X} \{\varphi^k(x) + [\varepsilon_k F(x_k) - \nabla \varphi(x^k)]^T x + \varepsilon_k u(x)\}$$

$$\Leftrightarrow \quad \min_{x\in X} \left\{ \frac{\varphi(x)}{\varepsilon_k} + \left[F(x^k) - \frac{\nabla\varphi(x^k)}{\varepsilon_k} \right]^T x + u(x) \right\}$$

on rebaptise $\frac{\varphi}{\varepsilon_k}$ par φ_k $\forall\ k \geq 0$:

$$\Leftrightarrow \quad \min_{x\in X} \left\{ \varphi^k(x) + [F(x^k) - \nabla\varphi^k(x^k)]^T x + u(x) \right\}$$

on rajoute des constantes :

$$\Leftrightarrow \quad \min_{x\in X} \left\{ \varphi^k(x) - \varphi^k(x^k) + u(x) - u(x^k) + [F(x^k) - \nabla\varphi^k(x^k)]^T (x - x^k) \right\}$$

$$\Leftrightarrow \quad (\mathrm{CNNLP}^k_{\varphi^k})$$

■

Cet algorithme de Cohen dont nous venons de détailler une itération va générer une suite d'itérés $\{x^k\}$ telle que x^k est solution du problème auxiliaire de Cohen pour tout k . Sous certaines hypothèses, cette suite $\{x^k\}$ converge vers x^* une solution de (AVIP). Ces hypothèses sont les suivantes :

- F est fortement monotone de module m_F et Lipschitz continu de module M_F sur X ;
- φ est fortement convexe de module m_φ et de gradient Lipschitz continu de module $M_{\nabla\varphi}$ sur X ;
- φ est telle que $2m_\varphi > \frac{M_F^2}{m_F}$.

Nous nous permettons d'insister tout particulièrement sur l'hypothèse de monotonicité forte de la fonction F car dans le prochain chapitre, nous tenterons de montrer qu'en ôtant cette hypothèse et en la remplaçant par l'hypothèse moins forte de co-coercivité de F , la convergence de l'algorithme de Cohen n'en sera pas modifiée.

3.7 Conclusions

Dans ce troisième chapitre, nous avons étudié les problèmes d'inéquation variationnelle non symétriques qui constituent la deuxième classe de cas particuliers de (GVIP). Ces problèmes sont caractérisés par le fait que la fonction F définissant l'inéquation variationnelle ne peut s'écrire sous la forme d'un gradient. Par conséquent, aucune fonction de mérite utilisable dans l'algorithme d'approximation du coût n'était connue.

Pour résoudre ce problème, nous avons introduit une classe de fonctions intéressantes que nous avons appelées *fonctions de mérite*. Nous avons tout d'abord étudié les propriétés de ces fonctions qui sont égales au *gap* dans l'inégalité de Fenchel de la fonction $u + \delta_X + \varphi$. Elles sont, de plus, continues (sous quelques conditions) et elles possèdent certaines caractéristiques des fonctions convexes comme la coercivité faible et des ensembles de niveau bornés. Grâce à ces fonctions, nous avons ensuite généré un algorithme de descente comprenant à la fois la recherche d'une direction de descente via la résolution d'un problème d'optimisation et ensuite une recherche linéaire exacte le long de cette direction relativement à la fonction de mérite. Nous avons établi la convergence de ce processus itératif sous un certain nombre d'hypothèses sur le problème dont la monotonicité forte de la fonction F. Finalement, nous avons proposé un algorithme basé sur le principe du problème auxiliaire et nous avons montré l'équivalence des sous-problèmes caractérisant chaque itération de chacun des algorithmes.

Il serait intéressant d'étendre les résultats présentés dans ce chapitre en établissant la convergence de l'algorithme de descente dans le cas où le sous-problème (3.2) de recherche de la direction de descente est résolu approximativement et où on admettrait une solution inexacte pour la recherche linéaire. Mais nous nous sommes limités au cas où l'algorithme était exact.

Jusqu'à présent, nous nous sommes contentés de citer le principe du problème auxiliaire. Nous avons admis sa convergence sous l'hypothèse de monotonicité forte. Une seconde source d'intérêt serait l'étude du processus itératif basé sur ce principe du problème auxiliaire et plus précisément la convergence de cet algorithme sous des hypothèses plus faibles que celle de monotonicité forte. Dans le chapitre suivant, nous suivrons cette piste en élargissant les hypothèses de convergence de l'algorithme de Cohen à la co-coercivité de la fonction F.

Chapitre 4

La co-coercivité et son rôle dans la convergence du processus itératif de Cohen pour la résolution des inéquations variationnelles

4.1 Introduction

Dans les chapitres précédents,nous avons étudié plusieurs algorithmes itératifs, l'algorithme d'approximation du coût, l'algorithme de descente et l'algorithme basé sur le principe du problème auxiliaire. Ceux-ci ont été introduits dans le but de résoudre des inéquations variationnelles. Nous avons montré que les deux premiers algorithmes étaient convergents moyennant certaines hypothèses et nous avons affirmé sans démonstration que l'algorithme basé sur le principe du problème auxiliaire l'était également. Nous avons insisté sur le fait que, parmi ces hypothèses, on retrouve celle selon laquelle la fonction F doit être fortement monotone, ce qui est une hypothèse assez restrictive.

Dans ce chapitre, nous allons introduire la propriété de co-coercivité d'une fonction et nous montrerons que l'algorithme basé sur le principe du problème auxiliaire associé à l'inéquation variationnelle exposée ci-dessous converge même sous l'hypothèse moins forte de co-coercivité de F . Finalement, nous généraliserons ce principe au cas où F n'est plus approximée par un gradient mais par une fonction quelconque. Nous étudierons la convergence de cet algorithme généralisé et nous en verrons une application.

Soient X un ensemble convexe fermé non vide de $I\!R^n$;

$F : X \to I\!R^n$ une fonction continue et monotone;

$u : X \to I\!R$ une fonction convexe, continue sur X .

Le problème à résoudre est exprimé sous la forme suivante :

(AVIP) | Trouver $x^* \in X$ et $\xi_u^* \in \partial u(x^*)$ tels que

$$[F(x^*) + \xi_u^*]^T(x - x^*) \geq 0 \qquad \forall\, x \in X\,.$$

Au chapitre 3, nous avons déjà introduit une formulation de (AVIP) plus facile à manipuler. À cet effet, nous avions ajouté une hypothèse de régularité (paragraphe 3.1) que nous supposons à nouveau vérifiée tout au long de ce chapitre.

Hypothèse 4.1
$\mathrm{int}(\mathrm{dom}\, u) \cap X \neq \emptyset$.

La formulation équivalente de (AVIP) est la suivante :

Trouver $x^* \in X$ tel que

$$F(x^*)^T(x - x^*) + u(x) - u(x^*) \geq 0 \qquad \forall\, x \in X\,. \tag{4.1}$$

Nous attirons votre attention sur le fait que, comme (AVIP) et (4.1) sont des problèmes équivalents, nous étudierons leurs propriétés sans les distinguer.

Nous supposons que les hypothèses sur X , F et u sont valables tout au long du chapitre. Nous admettons également que (AVIP) possède au moins une solution, i.e. que F est monotone sur X et

- soit $\mathrm{dom}\, u \cap X$ est borné ($\mathrm{int}(\mathrm{dom}\, u) \cap X$ étant non vide);
- soit la condition de coercivité suivante est vérifiée :

$$\exists\, \overline{x} \in \mathrm{dom}\, u \cap X \quad \text{tel que} \quad \lim_{\substack{x \in X \\ \|x\| \to \infty}} \left\{ \frac{F(x)^T(x - \overline{x}) + u(x)}{\|x\|} \right\} = +\infty\,.$$

L'ensemble de ces solutions est noté Ω .

4.2 Fonctions co-coercives

Nous allons, à présent, définir la notion de co-coercivité d'une fonction et montrer qu'il s'agit d'un concept intermédiaire entre la monotonicité et la monotonicité forte d'une fonction.

Définition 4.1
La fonction F est *co-coercive de module α_F sur X* si et seulement si il existe une constante $\alpha_F > 0$ telle que

$$(F(y) - F(x))^T(y - x) \geq \alpha_F \, \|F(y) - F(x)\|^2 \qquad \forall\, x, y \in X \,.$$

Proposition 4.1
Une fonction co-coercive est monotone.

Proposition 4.2
Une fonction co-coercive n'est pas nécessairement fortement monotone.

Afin d'illustrer cette propriété, nous proposons l'exemple suivant :
Soient $F : X \longrightarrow I\!R^n$ et $b \in I\!R^n$ tels que $\forall\, x \in X \;\; F(x) = b$.

- On vérifie facilement que F est co-coercive car

$$\exists\; \alpha_F > 0 \;\text{ tel que}$$
$$(F(y) - F(x))^T(y - x) \geq \alpha \, \|F(y) - F(x)\|^2 \;\; \forall\, x, y \in X \,.$$

- Par contre, F n'est pas fortement monotone (définition 1.12) car

$$\forall\; m_F > 0 \quad \exists\; x, y \in X \;\text{ tels que}$$
$$(F(y) - F(x))^T(y - x) < m_F \, \|y - x\|^2 \,.$$

  En effet, comme $F(y) = F(x) = b$, il suffit de choisir $x, y \in X$ tels que x et y soient distincts pour que l'inégalité suivante soit vérifiée :

$$0 < m_F \, \|y - x\|^2 \,.$$

Proposition 4.3
Une fonction fortement monotone et Lipschitz continue est co-coercive.

Rassemblant ces propriétés, il apparait clairement que la co-coercivité est une notion intermédiaire entre la monotonicité et la forte monotonicité.

Proposition 4.4
La somme de fonctions co-coercives est une fonction co-coercive.

Les démonstrations de ces propriétés sont détaillées en annexe A.7. Les deux suivantes sont admises sans preuve.

Proposition 4.5
Soit $F : X \rightarrow I\!R^n$ *une fonction inversible.*
Alors F *est co-coercive sur* X *ssi* F^{-1} *est Lipschitz continue et fortement monotone sur son image* $\mathrm{Im}\, F$.

Proposition 4.6
Une fonction co-coercive est monotone et lipschitzienne sur son domaine de définition.

4.3 Convergence du cadre itératif basé sur le principe du problème auxiliaire sous l'hypothèse de co-coercivité

Soit Ω l'ensemble des solutions de (AVIP). Rappelons que nous avons supposé, dans la section 4.1, qu'il était non vide.

4.3.1 Principe du problème auxiliaire

Soient $\varepsilon_k > 0$, $x^k \in X$ un itéré donné et Φ^k une fonction approximant F . Le problème auxiliaire d'inéquation variationnelle consiste à trouver $x^{k+1} \in X$ tel que

$$[\varepsilon_k F(x^k) + \Phi(x^{k+1}) - \Phi(x^k)]^T (x - x^{k+1}) + \varepsilon_k (u(x) - u(x^{k+1})) \geq 0 \quad \forall\, x \in X \,. \tag{4.2}$$

Nous avons déjà traité le problème auxiliaire de Cohen associé à (AVIP) au chapitre 3. Mais ce dernier était exprimé sous la forme d'un problème de minimisation.

Soient $\varphi^k : X \to I\!R$ une fonction fortement convexe et de classe C^1 sur X et $\varepsilon^k > 0$. Étant donné $x^k \in X$, trouver x^{k+1} tel que

$$x^{k+1} \in \underset{x \in X}{\operatorname{argmin}}\{\varphi^k(x) + [\varepsilon_k F(x^k) - \nabla\varphi^k(x^k)]^T x + \varepsilon_k u(x)\} . \tag{4.3}$$

Nous pouvons affirmer que ces deux formulations sont équivalentes si $\nabla\varphi^k \equiv \Phi^k$ pour tout k .

Preuve :
Soit $x^k \in X$ donné.

$$\begin{aligned}
(4.3) \quad \Leftrightarrow \quad & \text{Trouver } x^{k+1} \text{ tel que} \\
& x^{k+1} \in \underset{x \in X}{\operatorname{argmin}}\{\varphi^k(x) + [\varepsilon_k F(x^k) - \nabla\varphi^k(x^k)]^T x + \varepsilon_k u(x)\} \\
\Leftrightarrow \quad & \text{trouver } x^{k+1} \in X \text{ tel que} \\
& 0 \in \partial(\varphi^k(x) + [\varepsilon_k F(x^k) - \nabla\varphi^k(x^k)]^T x + \varepsilon_k u(x) + \delta_X(x))_{|x=x^{k+1}}
\end{aligned}$$

par la proposition 1.17

$$\begin{aligned}
\Leftrightarrow \quad & \text{Trouver } x^{k+1} \in X \text{ tel que} \\
& 0 \in \nabla\varphi^k(x^{k+1}) + \varepsilon_k F(x^k) - \nabla\varphi^k(x^k) + \varepsilon_k \partial u(x^{k+1}) + N_X(x^{k+1})
\end{aligned}$$

car les fonctions sont propres et convexes,
$\operatorname{rint}(\operatorname{dom}\varphi^k) \cap \operatorname{rint}(\operatorname{dom} u) \cap \operatorname{rint}(\operatorname{dom}\delta_X) \neq \emptyset$ et $\partial\delta_X(x) = N_X(x)$ (cfr[9, p. 223])

$$\begin{aligned}
\Leftrightarrow \quad & \text{Trouver } x^{k+1} \in X \text{ et } \xi_u^{k+1} \in \partial u(x^{k+1}) \text{ tels que} \\
& -\nabla\varphi^k(x^{k+1}) - \varepsilon_k F(x^k) + \nabla\varphi^k(x^k) - \varepsilon_k \xi_u^{k+1} \in N_X(x^{k+1}) \\
\Leftrightarrow \quad & \text{Trouver } x^{k+1} \in X \text{ et } \xi_u^{k+1} \in \partial u(x^{k+1}) \text{ tels que} \\
& (\Phi^k(x^{k+1}) + \varepsilon_k F(x^k) - \Phi^k(x^k) + \varepsilon_k \xi_u^{k+1})^T (x - x^{k+1}) \geq 0 \\
& \forall\, x \in X
\end{aligned}$$

car $\Phi^k \equiv \nabla\varphi^k \quad \forall\, k$

$$\Leftrightarrow \quad (4.2)$$

par [18], [19, Th. 27.4], [20], [21] et l'hypothèse (4.1). ∎

Le principe du problème auxiliaire peut être utilisé dans un algorithme dont (4.2) ou (4.3) constitue une itération. Afin d'alléger l'écriture, nous décidons d'appeler cet algorithme *processus itératif de Cohen.*

4.3.2 Convergence du processus itératif de Cohen

Aux chapitres 2 et 3, nous avons déjà rencontré le processus itératif de Cohen permettant de résoudre des inéquations variationnelles. Nous avons vu que cet algorithme générait une suite d'itérés $\{x^k\}$ appartenant à X et vérifiant (4.3) pour tout k . Nous avions admis que cette suite convergeait vers la solution x^* du problème d'inéquation variationnelle. Nous avions également insisté sur le fait que cette convergence nécessitait l'hypothèse restrictive de monotonicité forte de la fonction F .

Comme annoncé dans l'introduction de ce chapitre, nous désirons relâcher cette hypothèse en la remplaçant par celle de co-coercivité de F . Nous allons montrer dans le théorème suivant que le processus itératif de Cohen converge sous l'hypothèse plus faible de co-coercivité.

Pour démontrer le théorème de convergence suivant, nous allons utiliser les propositions 1.3 et 1.7 qui caractérisent les fonctions fortement convexes et de gradient Lipschitz continu.

Théorème 4.1 (Convergence du processus itératif de Cohen)
Soient F une fonction co-coercive sur X de module γ_F et $\Phi = \nabla\varphi$. Soit φ une fonction fortement convexe sur X de module m_φ et $\nabla\varphi$ lipschitzien sur X de module $M_{\nabla\varphi}$.
Alors

1. *il existe une solution unique x^{k+1} de (4.2).*

Si, de plus, nous choisissons $\varepsilon = \varepsilon_k \ \forall\ k$ tel que $0 < \varepsilon < 2m_\varphi\gamma_F$,
alors

2. *la suite des itérés $\{x^k\}$ est bornée;*

3. *la suite $\{x^k\}$ converge vers une solution de (AVIP).*

Preuve :

1. Comme φ est fortement convexe sur X , la solution x^{k+1} de (4.3) est unique par la proposition 1.19. Comme (4.2) est équivalent à (4.3), la première partie de la thèse est prouvée.

2. **Montrons que $\{x^k\}$ est bornée :**

Soit x^* une solution de (AVIP).

On définit la fonction

$$\Lambda^*(x) = \varphi(x^*) - \varphi(x) - \nabla\varphi(x)^T(x^* - x) \quad \forall\, x \in X\,.$$

Comme φ est fortement monotone, par la proposition 1.3,

$$\Lambda^*(x) \geq \frac{m_\varphi}{2}\, \|x - x^*\|^2 \geq 0 \qquad \forall\, x \in X\,. \tag{4.4}$$

Montrons tout d'abord que $\{\Lambda^*(x^k)\}$ est une suite décroissante.

Soit x^{k+1} la solution de (4.2) ,

$$\begin{aligned} \Lambda^*(x^k) - \Lambda^*(x^{k+1}) &= \varphi(x^{k+1}) - \varphi(x^k) - \nabla\varphi(x^k)^T(x^{k+1} - x^k) \\ &\quad + [\nabla\varphi(x^{k+1}) - \nabla\varphi(x^k)]^T(x^* - x^{k+1})\,. \end{aligned} \tag{4.5}$$

Comme

$$[\varepsilon F(x^k) + \nabla\varphi(x^{k+1}) - \nabla\varphi(x^k)]^T(x - x^{k+1}) \geq \varepsilon(u(x^{k+1}) - u(x)) \quad \forall\, x \in X$$

puisque x^{k+1} est solution de (4.2), nous pouvons déduire que cette inégalité est en particulier vraie pour $x^* \in X$.

Ainsi,

$$[\nabla\varphi(x^{k+1}) - \nabla\varphi(x^k)]^T(x^* - x^{k+1}) \geq \varepsilon(u(x^{k+1}) - u(x^*)) + \varepsilon F(x^k)^T(x^{k+1} - x^*)\,. \tag{4.6}$$

En rassemblant (4.5) et (4.6) et en utilisant la convexité forte de φ , nous avons

$$\begin{aligned} \Lambda^*(x^k) - \Lambda^*(x^{k+1}) &\geq \frac{m_\varphi}{2}\, \|x^k - x^{k+1}\|^2 + \varepsilon(u(x^{k+1}) - u(x^*)) \\ &\quad + \varepsilon F(x^k)^T(x^{k+1} - x^*)\,. \end{aligned} \tag{4.7}$$

En posant $x = x^{k+1}$ dans (AVIP), nous obtenons

$$u(x^{k+1}) - u(x^*) \geq -F(x^*)^T(x^{k+1} - x^*)\,. \tag{4.8}$$

Combinant (4.7) et (4.8), nous pouvons écrire

$$\begin{aligned} \Lambda^*(x^k) - \Lambda^*(x^{k+1}) &\geq \frac{m_\varphi}{2}\, \|x^k - x^{k+1}\|^2 + \varepsilon[F(x^k) - F(x^*)]^T(x^{k+1} - x^*) \\ &\geq \frac{m_\varphi}{2}\, \|x^k - x^{k+1}\|^2 + T \end{aligned} \tag{4.9}$$

où

$$T = \varepsilon[F(x^k) - F(x^*)]^T(x^{k+1} - x^*) \; .$$

Comme F est co-coercive de constante γ_F , nous avons

$$\begin{aligned} T &\geq \varepsilon[\gamma_F \; \|F(x^k) - F(x^*)\|^2 + [F(x^k) - F(x^*)]^T(x^{k+1} - x^k)] \\ &\geq -\frac{\varepsilon}{4\gamma_F} \; \|x^{k+1} - x^k\|^2 \; . \end{aligned} \tag{4.10}$$

La dernière inégalité s'explique par le fait que

$$\|2\gamma_F \; [(F(x^k) - F(x^*)] + (x^{k+1} - x^k)\|^2 \geq 0 \; .$$

Donc, en rassemblant (4.10) et (4.9), nous pouvons affirmer que

$$\Lambda^*(x^k) - \Lambda^*(x^{k+1}) \geq \frac{1}{2}\left(m_\varphi - \frac{\varepsilon}{2\gamma_F}\right) \|x^k - x^{k+1}\|^2 \; . \tag{4.11}$$

Comme, par hypothèse, ε est tel que $0 < \varepsilon < 2m_\varphi\gamma_F$, nous avons $\Lambda^*(x^k) \geq \Lambda^*(x^{k+1})$, c'est-à-dire que la suite $\{\Lambda^*(x^k)\}$ est décroissante.

Nous devons, à présent, distinguer deux cas :

- Soit $x^{k+1} = x^k$.
  Alors, comme x^{k+1} est solution de (4.2), nous avons

  $$0 \leq \varepsilon_k[F(x^k)]^T(x - x^k) + \varepsilon_k(u(x) - u(x^k)) \qquad \forall \; x \in X \; .$$

  Comme $\varepsilon_k = \varepsilon > 0 \; \forall \; k$ par hypothèse,

  $$F(x^k)^T(x - x^k) - u(x^k) + u(x) \geq 0 \qquad \forall \; x \in X \; .$$

  Et donc x^k est solution de (4.1) et de (AVIP).

- Soit $x^{k+1} \neq x^k$.
  Alors $\{\Lambda^*(x^k)\}$ est strictement décroissante.
  Or, comme $\{\Lambda^*(x^k)\}$ est minorée par zéro (par (4.4)), elle est convergente vers sa borne inférieure.
  Comme toute suite convergente est de Cauchy, $\{\Lambda^*(x^k)\}$ est de Cauchy.
  Par (4.11), $\{x^k\}$ est également de Cauchy.
  Or, comme toute suite de Cauchy est bornée, $\{x^k\}$ est bornée.

3. **Montrons que $\{x^k\}$ converge vers une solution de (AVIP) :**

Rappelons que $\{x^k\}$ est la suite engendrée par le processus itératif de Cohen. Soit $\{x^k\}_{k\in K'}$ une sous-suite de $\{x^k\}$, $K' \subset K$ où K est l'index de la suite engendrée par l'algorithme de Cohen. Cette sous-suite vérifie (4.2) pour tout k .

Soit $\overline{x}$ la limite de $\{x^k\}_{k\in K'}$. En passant à la limite dans (4.2), nous obtenons que $\overline{x}$ est solution de (AVIP).

Nous pouvons remplacer x^* par $\overline{x}$ dans tous les résultats prouvés précédemment puisque x^* était une solution quelconque de (AVIP).

Nous définissons

$$\overline{\Lambda}(x) \triangleq \varphi(\overline{x}) - \varphi(x) - \nabla\varphi(x)^T(\overline{x} - x) \qquad \forall\, x \in X\,.$$

Nous pouvons affirmer que $\{\overline{\Lambda}(x^k)\}_{k\in K'}$ est décroissante et

$$\overline{\Lambda}(x^k) \geq \frac{m_\varphi}{2}\ \|x^k - \overline{x}\|^2 \qquad \forall\, k \in K'\,.$$

Comme, par hypothèse, $\nabla\varphi$ est lipschitzien, nous avons

$$\overline{\Lambda}(x^k) \leq \frac{1}{2}M_{\nabla\varphi}\ \|\overline{x} - x\|^2 \qquad \forall\, k \in K'\,.$$

Ainsi,

$$\frac{m_\varphi}{2}\ \|x^k - \overline{x}\|^2 \leq \overline{\Lambda}(x^k) \leq \frac{M_{\nabla\varphi}}{2}\ \|\overline{x} - x^k\|^2 \qquad \forall\, k \in K'\,.$$

Comme $\{x^k\}_{k\in K'} \underset{k\to\infty}{\longrightarrow} \overline{x}$, $\{\overline{\Lambda}(x^k)\}_{k\in K'} \underset{k\to\infty}{\longrightarrow} 0$.

Or, si une sous-suite d'une suite monotone converge vers zéro, alors la suite converge vers 0 .

Donc, $\{\overline{\Lambda}(x^k)\} \underset{\substack{k\to\infty\\ k\in K}}{\longrightarrow} 0$.

Et dès lors, $\{x^k\}_{k\in K} \underset{k\to\infty}{\longrightarrow} \overline{x}$ car $\frac{m_\varphi}{2}\ \|x^k - \overline{x}\|^2 \leq \overline{\Lambda}(x^k)\ \forall\, k \in K$. ∎

Ce théorème est le résultat clé du chapitre car il généralise la convergence du processus itératif de Cohen aux fonctions co-coercives et pas nécessairement fortement monotones.

4.4 Généralisation du principe du problème auxiliaire

Le principe du problème auxiliaire étudié jusqu'à présent était caractérisé par le fait que la fonction F était toujours approximée par le gradient d'une fonction. Dans le processus itératif de Cohen dont une itération est développée en (4.3), F était approximée par $\nabla\varphi^k$. L'approximation changeait à chaque itération. Dans ce paragraphe, nous allons généraliser le processus itératif de Cohen en introduisant une nouvelle fonction d'approximation.

4.4.1 Processus itératif de Cohen généralisé

Une itération du processus itératif de Cohen généralisé consiste en la résolution du problème exposé ci-dessous :

Étant donné $x^k \in X$ et $\varepsilon > 0$,
trouver $x^{k+1} \in X$ tel que $\forall\ x \in X$

$$[\varepsilon F(x^k) + \Gamma_\varepsilon(x^{k+1}) - \Gamma_\varepsilon(x^k)]^T(x - x^{k+1}) + \varepsilon(u(x) - u(x^{k+1})) \geq 0 \tag{4.12}$$

avec

$$\Gamma_\varepsilon(x) = \varepsilon\Gamma_1(x) + \Gamma_2(x) \qquad \forall\ x \in X$$

où Γ_1 et Γ_2 sont des fonctions non linéaires définies de X dans $I\!R^n$.

On peut, par exemple, choisir

$$\Gamma_1(x) = Mx + L(x)\ , \quad \Gamma_2(x) = \nabla\varphi(x) \qquad \forall\ x \in X$$

où M est une matrice symétrique, L une fonction monotone non linéaire et φ une fonction convexe de gradient lipschitzien.

Si nous choisissons $\Gamma_1 \equiv 0$, nous retrouvons le principe du problème auxiliaire non généralisé. En effet,

$$\Gamma_\varepsilon(x) \;=\; \varepsilon\Gamma_1(x) + \Gamma_2(x)$$

se réduit à

$$\Gamma_\varepsilon(x) \;=\; \Gamma_2(x) \;=\; \nabla\varphi(x)\ .$$

Nous pouvons à présent établir un résultat de convergence également basé sur la notion de co-coercivité. La preuve de ce théorème étant fort semblable à celle du théorème

4.1 de convergence du processus itératif de Cohen non généralisé, nous jugeons bon de ne pas la développer.

Théorème 4.2 (Convergence du processus itératif de Cohen généralisé)
Soient φ une fonction convexe, $\nabla\varphi$ son gradient fortement monotone de module $m_{\nabla\varphi}$ et $F-L$ une fonction co-coercive de module γ_{F-L} sur X .
Soit $\varepsilon > 0$ tel que

$$\begin{cases} 0 < \varepsilon \leq \dfrac{2m_{\nabla\varphi}\gamma_{F-L}}{1-2\gamma_{F-L}\lambda_{\min}(M)} & si \ \dfrac{1}{2\gamma_{F-L}} - \lambda_{\min}(M) > 0 \\ \\ 0 < \varepsilon < \infty & sinon \end{cases}$$

avec $\lambda_{\min}(M)$ définie comme la plus petite valeur propre de la matrice M .
Alors,

1. *il existe une unique solution à (4.12);*
2. *la suite $\{x^k\}$ des itérés obtenus par le processus de Cohen généralisé est bornée;*
3. *la suite $\{x^k\}$ converge vers une solution de (AVIP).*

4.4.2 Application

La formulation générale du problème auxiliaire de Cohen possède de nombreuses applications. Dans ce paragraphe, nous montrons que le problème d'inéquation variationnelle proximale est un cas particulier du problème auxiliaire généralisé.

Soit F la fonction monotone utilisée pour définir l'inéquation variationnelle (AVIP) que nous avons introduite au début du chapitre. Nous décomposons cette fonction F en la somme d'une fonction co-coercive F_1 et d'une fonction monotone F_2 . Nous choisissons $\Gamma_1 = F_2$ dans le problème auxiliaire de Cohen généralisé.
En transformant l'écriture du problème (4.12) dans ce cas particulier, nous obtenons le problème suivant :

Étant donné $x^k \in X$ et $\varepsilon > 0$, trouver $x^{k+1} \in X$ tel que

$$\begin{aligned} [\varepsilon(F_1(x^k) + F_2(x^{k+1})) + \Gamma_2(x^{k+1}) - \Gamma_2(x^k)]^T(x - x^{k+1}) & \\ +\varepsilon(u(x) - u(x^{k+1})) & \geq 0 \end{aligned} \tag{4.13}$$

pour tout $x \in X$.

Comme toute fonction monotone F peut s'exprimer comme la somme de la fonction nulle (trivialement co-coercive) et d'elle-même, nous retrouvons comme cas particulier de (4.13) l'inéquation variationnelle proximale donnée par

$$[\varepsilon F(x^{k+1}) + \Gamma_2(x^{k+1}) - \Gamma_2(x^k)]^T(x - x^{k+1}) + \varepsilon(u(x) - u(x^{k+1})) \geq 0$$

pour tout $x \in X$.

4.5 Conclusions

Dans ce dernier chapitre, nous avons porté un intérêt particulier à l'étude de l'algorithme engendré par le principe du problème auxiliaire. Étant donné l'inéquation variationnelle suivante à résoudre :

Trouver $x^* \in X$ tel que
$$F(x^*)^T(x - x^*) + u(x) - u(x^*) \geq 0 \qquad \forall\, x \in X\ ,$$

avec F une fonction co-coercive (et pas fortement monotone), nous avons démontré que le processus itératif de Cohen convergeait vers une solution de cette équation (en supposant qu'il en existe une). De cette façon, nous avons réduit les conditions de convergence du processus de Cohen. En effet, la convergence ne nécessite plus la monotonicité forte mais la co-coercivité de la fonction F .

D'autre part, nous avons généralisé l'algorithme de Cohen au sens où la correspondance approximant F n'est plus le gradient d'une fonction quelle qu'elle soit, comme dans les chapitres précédents, mais une fonction non linéaire quelconque Γ_ε .

Nous avons finalement tenté de mettre en évidence l'utilité de cette généralisation à travers une application connue : l'inéquation variationnelle proximale.

Une fois de plus, nous n'avons pas su présenter dans ce chapitre tous les concepts qui nous semblaient intéressants. Nous pouvons notamment mentionner les notions de fonctions partiellement co-coercives et de systèmes d'inéquations variationnelles.

Conclusions générales

Nous avons donc construit, avec M. PATRIKSSON, un cadre algorithmique qui contient trois algorithmes : l'algorithme d'approximation du coût, l'algorithme de descente, le processus itératif de Cohen. Tout au long du mémoire, nous avons étudié diverses propriétés de ces algorithmes : ceux-ci sont basés sur la résolution d'un sous-problème obtenu par approximation de la fonction de coût originale; ces problèmes génèrent tous des directions de descente; les solutions du problème initial sont aussi solutions du sous-problème. De plus, les algorithmes convergent tous vers une solution du problème d'inéquation variationnelle.

Dans notre étude, les algorithmes d'approximation du coût et de Cohen ont été employés pour résoudre des problèmes symétriques de minimisation d'une fonction (dans le chapitre 2, la fonction était $T = f + u$). Nous aurions pu proposer une extension de ceux-ci pour permettre la résolution du problème qui consiste à trouver un zéro d'un opérateur monotone maximal quelconque. A RENAUD et G. COHEN ont déjà présenté une généralisation du processus itératif de Cohen. Un sujet de recherche intéressant serait à présent d'adapter l'agorithme d'approximation du coût à ce nouveau problème.

Annexes

Annexe A.1 Équivalence des différentes formulations de (GVIP)

Soient$F : X \to I\!R^n$ et $U : X \to 2^{I\!R^n}$ deux opérateurs, X un ensemble convexe fermé non vide dans $I\!R^n$.

Montrons que (GVIP) est équivalent au problème suivant :

$$\begin{aligned}&\text{Trouver } x^* \in X \text{ tel que}\\ &0 \in F(x^*) + U(x^*) \in N_X(x^*)\ .\end{aligned}$$

Preuve :

$$\begin{aligned}\text{(GVIP)} \quad \Leftrightarrow \quad & \text{Trouver } x^* \in X \text{ et } u^* \in U(x^*) \text{ tels que}\\ & [F(x^*) + u^*]^T(x - x^*) \geq 0 \qquad \forall\, x \in X\\ \Leftrightarrow \quad & \text{Trouver } x^* \in X \text{ et } u^* \in U(x^*) \text{ tels que}\\ & -(F(x^*) + u^*) \in N_X(x^*)\\ \Leftrightarrow \quad & \text{Trouver } x^* \in X \text{ tel que}\\ & 0 \in N_X(x^*) + F(x^*) + U(x^*)\ .\end{aligned}$$

■

Annexe A.2 Lipschitz continuité de ∇f

Nous allons à présent rappeler et démontrer la proposition 1.7 :

Soit X un ensemble convexe fermé non vide de $I\!R^n$.
Soit $f : X \to I\!R^n$ une fonction de classe C^1 sur X .
Si ∇f est Lipschitz continu de module $M_{\nabla f}$ sur X , alors

$$\exists\ M_{\nabla f} \geq 0 \ \textit{tel que}\ \forall\ x, y \in X$$
$$f(y) - f(x) \leq \nabla f(x)^T (y - x) + \tfrac{M_{\nabla f}}{2}\ \|x - y\|^2 .$$

Preuve :
Nous supposons que ∇f est Lipschitz continu, i.e. que

$$\exists\ M_{\nabla f} \geq 0 \ \text{tel que}\ \forall\ x, y \in X$$
$$\|\nabla f(x) - \nabla f(y)\| \leq M_{\nabla f}\ \|x - y\| .$$

Nous définissons la fonction

$$\Phi(t) = f((1-t)x + ty) - f(x) - t[\nabla f(x)]^T (y - x) \quad \forall\ x, y \in X\ ,\ \forall\ t \in [0,1] .$$

Cette fonction est telle que

$$\begin{aligned} \nabla\Phi(t) &= [\nabla f((1-t)x + ty)]^T (y - x) - \nabla f(x)^T (y - x) \\ &= [\nabla f((1-t)x + ty - \nabla f(x)]^T (y - x) \end{aligned}$$

et

$$\begin{aligned} \Phi(1) &= f(y) - f(x) - \nabla f(x)^T (y - x) \qquad \forall\ x, y \in X \\ &= \Phi(0) + \int_0^1 \nabla\Phi(t)\ dt \\ &= \int_0^1 \nabla\Phi(t)\ dt \end{aligned}$$

car $\Phi(0) = 0$

$$\begin{aligned} &\leq \int_0^1 |\nabla\Phi(t)|\ dt \\ &\leq \int_0^1 |[\nabla f((1-t)x + ty) - \nabla f(x)]^T (y - x)|\ dt \end{aligned}$$

$$
\begin{aligned}
&\leq \int_0^1 \|\nabla f((1-t)x+ty)-\nabla f(x)\|\ \|y-x\|\ dt \\
&\leq \int_0^1 \|t(y-x)\|\ M_{\nabla f}\ \|y-x\|\ dt
\end{aligned}
$$

car ∇f est Lipschitz continu

$$
\begin{aligned}
&\leq M_{\nabla f}\ \|y-x\|^2 \int_0^1 t\ dt \\
&\leq \frac{M_{\nabla f}}{2}\ \|y-x\|^2 .
\end{aligned}
$$

Donc,

$$f(y)-f(x) \leq \nabla f(x)^T(y-x)+\frac{M_{\nabla f}}{2}\ \|y-x\|^2 \qquad \forall\, x,y \in X .$$

■

Annexe A.3 Condition d'unicité d'un minimum

Nous démontrons dans cette annexe une proposition qui a été énoncée au premier chapitre concernant les préliminaires.

Proposition 1.19

Soit $f : X \to I\!R \cup \{-\infty, +\infty\}$ une fonction fortement convexe de constante m_f , propre et semi-continue inférieurement.

Alors f possède un unique minimum x_0 et celui-ci vérifie

$$f(x_0) \leq f(x) - \frac{m_f}{2}\|x - x_0\|^2 \qquad \forall\, x \in X\ .$$

Preuve :

Par [22, Prop. 6], nous savons que, si $0 \in \partial f(x_0)$, alors

$$f(x) \geq f(x_0) + \frac{m_f}{2}\ \|x - x_0\|^2 \qquad \forall\, x \in X\ .$$

Montrons, en utilisant cet argument que le minimum de f est unique.

Supposons par l'absurde qu'il existe deux minima distincts de f , x_1 et x_2 dans X , i.e. $0 \in \partial f(x_1)$ et $0 \in \partial f(x_2)$.

En utilisant l'argument cité ci-dessus, nous avons

$$f(x) \ \geq \ f(x_1) + \frac{m_f}{2}\ \|x - x_1\|^2 \qquad \forall\, x \in X$$

et

$$f(x) \ \geq \ f(x_2) + \frac{m_f}{2}\ \|x - x_2\|^2 \qquad \forall\, x \in X\ .$$

Donc

$$f(x_2) \geq f(x_1) + \frac{m_f}{2}\ \|x_2 - x_1\|^2\ .$$

Or, si $x_1 \neq x_2$, alors $\|x_2 - x_1\|^2 > 0$.

Donc, $f(x_2) > f(x_1)$.

Dès lors x_2 n'est pas un minimum de f . D'où la contraction attendue. Ainsi, f possède un unique minimum. ∎

Annexe A.4 Convexité de la fonction T_φ

Soient $\varphi : \mathbb{R}^n \to \mathbb{R}$ une fonction convexe, de classe C^1 sur $\operatorname{dom} u$;

$u : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ une fonction convexe, propre, s.c.i.;

$\operatorname{dom} u \subseteq \mathbb{R}^n$.

Nous avons défini

$$T_\varphi(x) = \varphi(x) + u(x) + f(y) - \varphi(y) + [\nabla f(y) - \nabla \varphi(y)]^T (x - y) \quad \forall\, x \in \mathbb{R}^n .$$

Voyons que, si φ ou u est strictement convexe sur $\operatorname{dom} u$, alors $T_\varphi(x)$ est strictement convexe sur $\operatorname{dom} u$, i.e.

$$\begin{aligned} &\forall\, x, z \in \operatorname{dom} u , \quad \forall\, \lambda \in [0,1] \\ &T_\varphi(\lambda x + (1-\lambda) z) \leq T_\varphi(x) + (1-\lambda) T_\varphi(z) . \end{aligned}$$

Preuve :
Soient $x, y \in$ $\operatorname{dom} u$ et $\lambda \in [0,1]$ quelconques.

$$\begin{aligned} T_\varphi(\lambda x + (1-\lambda)z) \;&=\; \varphi(x\lambda + (1-\lambda)z) + u(\lambda x + (1-\lambda)z) + f(y) - \varphi(y) \\ &\quad + [\nabla f(y) - \nabla \varphi(y)]^T (\lambda x + (1-\lambda) z - y) . \end{aligned}$$

Comme u ou φ est strictement convexe,

$$\begin{aligned} &T_\varphi(\lambda x + (1-\lambda)z) \\ &\quad < \;\lambda\varphi(x) + \lambda u(x) + (1-\lambda)\varphi(z) + (1-\lambda)u(z) + f(y) - \varphi(y) \\ &\qquad\quad + [\nabla f(y) - \nabla\varphi(y)]^T(\lambda x - y) + [\nabla f(y) - \nabla\varphi(y)]^T (1-\lambda) z \\ &\qquad\quad + \lambda f(y) - \lambda\varphi(y) - \lambda f(y) + \lambda\varphi(y) \\ &\quad < \;\lambda\varphi(x) + \lambda u(x) + \lambda f(y) - \lambda\varphi(y) \\ &\qquad\quad + [\nabla f(y) - \nabla\varphi(y)]^T(\lambda x - \lambda y + \lambda y - y) \\ &\qquad\quad + (1-\lambda)\varphi(z) + (1-\lambda)u(z) - \lambda f(y) + f(y) \\ &\qquad\quad - \varphi(y) + \lambda\varphi(y) + [\nabla f(y) - \nabla\varphi(y)]^T(1-\lambda) z \\ &\quad < \;\lambda T_\varphi(x) + (1-\lambda) T_\varphi(z) . \end{aligned}$$

■

On peut faire le même raisonnement avec u ou φ fortement convexe sur $\operatorname{dom} u$. Dans ce cas, on obtient $T_\varphi(\cdot)$ fortement convexe sur $\operatorname{dom} u$.

Annexe A.5 Équivalence des formulations du problème (3.2)

Soit $\varphi : I\!R^n \to I\!R \cup \{+\infty\}$ convexe, s.c.i. et $\varphi \in C^1$ sur X . Soit aussi

$$\begin{array}{rcl} L: \ X \times X & \to & I\!R \\ (x,y) & \rightsquigarrow & L(x,y) = u(x) - u(y) + \varphi(x) - \varphi(y) + [F(x) - \nabla\varphi(x)]^T(x-y) \ . \end{array}$$

Nous définissons $\psi : X \to I\!R^+ \cup \{+\infty\}$ comme

$$\boxed{\psi(x) = \sup_{y \in X} L(x,y) \ .} \tag{3.2}$$

Montrons l'équivalence du problème (3.2) avec les deux formulations suivantes :

1. **Première formulation** :

$$(3.2) \Leftrightarrow (3.4)$$

où (3.4) est le problème défini par

Soit $x \in X$. Trouver $y(x) \in X$ tel que

$$-(F(x) - \nabla\varphi(x)) \in \nabla\varphi(y(x) + \partial u(y(x)) + N_X(y(x)) \ .$$

Preuve :

$$\begin{array}{rcl} (3.2) & \Leftrightarrow & \text{Trouver } y(x) \in X \text{ tel que} \\ & & y(x) \text{ donne le } \sup\limits_{y \in X} L(x,y) \\ & \Leftrightarrow & \text{Trouver } y(x) \in X \text{ tel que} \\ & & y(x) \text{ donne le } \inf\limits_{y \in X} -L(x,y) \\ & \Leftrightarrow & \text{Trouver } y(x) \in X \text{ tel que} \\ & & 0 \in \partial(L(x,\cdot) + \delta_X)(y(x)) \\ & \Leftrightarrow & \text{Trouver } y(x) \in X \text{ tel que} \\ & & 0 \in \partial u(y(x)) + \partial\varphi(y(x)) + \partial\delta_X(y(x)) + F(x) - \nabla\varphi(x) \end{array}$$

car $\operatorname{rint} X \cap \operatorname{rint}(\operatorname{dom} u) \cap \operatorname{rint}(\operatorname{dom} \varphi) \neq \emptyset$ et

φ , u et δ_X sont des fonctions propres et convexes

$$\Leftrightarrow \text{ Trouver } y(x) \in X \text{ tel que}$$
$$0 \in \partial u(y(x)) + N_X(y(x)) + \nabla\varphi(y(x)) + F(x) - \nabla\varphi(x)$$

car $\varphi \in C^1$ sur X et $N_X(y(x)) = \partial\delta(y(x))$

$$\Leftrightarrow \ (3.4)$$

■

2. **Deuxième formulation :**

$$(3.2) \Leftrightarrow (3.3)$$

où (3.3) est le problème défini par

Soit $x \in X$. Trouver $y(x)$ tel que

$$[\nabla\varphi(y(x)) + F(x) - \nabla\varphi(x)]^T(y - y(x)) + u(y) - u(y(x)) \geq 0$$

$\forall\, y \in X$.

Preuve :

$$(3.2) \Leftrightarrow \text{ Trouver } y(x) \text{ tel que}$$
$$-(F(x) - \nabla\varphi(x)) \in \nabla\varphi(y(x)) + \partial u(y(x)) + N_X(y(x))$$

par 1.

$$\Leftrightarrow \text{ Trouver } y(x) \text{ et } \xi_u(y(x)) \in \partial u(y(x)) \text{ tels que}$$
$$-F(x) + \nabla\varphi(x) - \nabla\varphi(y(x)) - \xi_u(y) \in N_X(y(x))$$

$$\Leftrightarrow \text{ Trouver } y(x) \text{ et } \xi_u(y(x)) \in \partial u(y(x)) \text{ tels que}$$
$$[-F(x) + \nabla\varphi(x) - \nabla\varphi(y(x)) - \xi_u(y(x))]^T(y - y(x)) \leq 0$$
$$\forall\, y \in X$$

par [18], [20], [21], [19, Th. 27.4] et l'hypothèse 3.1

$$\Leftrightarrow \text{ Trouver } y(x) \text{ tel que}$$
$$[F(x) - \nabla\varphi(x) + \nabla\varphi(y(x))]^T(y - y(x)) + u(y) - u(y(x)) \geq 0$$
$$\forall\, y \in X .$$

■

Annexe A.6 Concavité de la fonction L

Soit $\varphi : I\!R^n \to I\!R \cup \{+\infty\}$ une fonction convexe, s.c.i. et de classe C^1 sur X .
Soit $u : I\!R^n \to I\!R \cup \{+\infty\}$ une fonction convexe, propre et s.c.i.
Soit X un ensemble convexe fermé non vide de $I\!R^n$.

Nous avons défini

$$\begin{array}{rcl} L: \ X \times X & \to & I\!R \\ (x,y) & \rightsquigarrow & L(x,y) = u(x) - u(y) + \varphi(x) - \varphi(y) + [F(x) - \nabla\varphi(x)]^T(x-y) \ . \end{array}$$

$L(x,\cdot)$ est concave par rapport à y sur X , c'est-à-dire

$$\forall\ \lambda \in [0,1]\ , \quad \forall\ z,y \in X$$
$$L(x, \lambda y + (1-\lambda)z) \geq \lambda L(x,y) + (1-\lambda)L(x,z)\ .$$

Preuve :

$$\begin{aligned} & L(x, \lambda y + (1-\lambda)z) \\ &= u(x) - u(\lambda y + (1-\lambda)z) + \varphi(x) - \varphi(\lambda y + (1-\lambda)z) \\ &\quad +[F(x) - \nabla\varphi(x)]^T(x - \lambda y - (1-\lambda)z)\ . \end{aligned}$$

Comme u et φ sont convexes,

$$\begin{aligned} & L(x, \lambda y + (1-\lambda)z) \\ &\geq u(x) - \lambda u(y) - (1-\lambda)u(z) + \varphi(x) - \lambda\varphi(y) - (1-\lambda)\varphi(z) \\ &\quad +[F(x) - \nabla\varphi(x)]^T(x - \lambda y + \lambda z - z) \\ &\geq u(x) + \varphi(x) + \lambda\varphi(x) + \lambda u(x) - \lambda u(x) - \lambda\varphi(x) - \lambda u(y) - \lambda\varphi(y) \\ &\quad +[F(x) - \nabla\varphi(x)]^T(\lambda x - \lambda y + x - \lambda x - z + \lambda z) \\ &\quad -(1-\lambda)u(z) - (1-\lambda)\varphi(z) \\ &\geq \lambda\varphi(x) + \lambda u(x) - \lambda u(y) - \lambda\varphi(y) \\ &\quad +[F(x) - \nabla\varphi(x)]^T\lambda(x-y) + (1-\lambda)u(x) + (1-\lambda)\varphi(x) \\ &\quad -(1-\lambda)u(z) - (1-\lambda)\varphi(z) + [F(x) - \nabla\varphi(x)]^T(1-\lambda)(x-z) \end{aligned}$$

$$\geq \lambda L(x,y) + (1-\lambda)L(x,z) \ .$$

∎

Si φ ou u est strictement convexe, alors $L(x,\cdot)$ est strictement concave par rapport à y .

De même, si φ ou u est fortement convexe, alors $L(x,\cdot)$ est fortement concave par rapport à y .

Annexe A.7 Co-coercivité

Soit $F : X \rightarrow I\!R^n$ où X est un ensemble convexe fermé non vide de $I\!R^n$ et F est une fonction continue.

Propriété 1
Si F est co-coercive sur X , alors F est monotone sur X .

Preuve :
Montrons que

$$\forall\; x, y \in X \qquad (F(y) - F(x))^T(y - x) \geq 0$$

et dans ce cas, par la définition 1.10, F sera monotone sur X .
On sait que F est co-coercive si et seulement si

$$\exists\; \alpha_F > 0 \;\text{ tel que }\; \forall\; x, y \in X$$
$$(F(y) - F(x))^T(y - x) \geq \alpha_F\; \|F(y) - F(x)\|^2\;.$$

Comme $\alpha_F > 0$ et $\|F(y) - F(x)\|^2 \geq 0$, nous obtenons la thèse. ■

Propriété 2
Si F est fortement monotone et Lipschitz continue sur X ,
alors F est co-coercive sur X .

Preuve :
Montrons qu'il existe $\alpha_F > 0$ tel que

$$\forall\; x, y \in X \qquad (F(y) - F(x))^T(y - x) \geq \alpha_F\; \|F(y) - F(x)\|^2\;.$$

Dans ce cas, par la définition 4.1, F sera co-coercive;
Nous savons que

$$\exists\; m_F > 0 \;\text{ tel que }\; \forall\; x, y \in X$$
$$(F(y) - F(x))^T(y - x) \geq m_F\; \|y - x\|^2$$

puisque F est fortement monotone (définition 1.12).

De plus,

$$\exists\ M_F > 0 \ \text{ tel que } \ \forall\ x, y \in X$$
$$\|F(y) - F(x)\| \leq M_F\ \|y - x\|$$

puisque F est lipschitzienne (définition 1.13).
Donc, en rassemblant ces inégalités, nous obtenons

$$\exists\ \alpha_F = \tfrac{m_F}{M_F^2} \ \text{ tel que}$$
$$(F(y) - F(x))^T(y - x) \geq \alpha_F\ \|F(y) - F(x)\|^2\ .$$

■

Propriété 3
Soient $F_i : X \longrightarrow I\!R^n$, $i = 1, \cdots, n$, *n fonctions co-coercives de constantes* α_{F_i} .
Alors $\sum\limits_{i=1}^{n} F_i \equiv F : X \longrightarrow I\!R^n$ *est co-coercive de constante* $\alpha_F = \min\limits_{i\in\{1,\cdots,n\}} \alpha_{F_i}$.

Preuve :
Montrons qu'il existe $\alpha_F > 0$ tel que $\forall\ x, y \in X$

$$(F(y) - F(x))^T(y - x) \geq \alpha_F\ \|F(y) - F(x)\|^2\ .$$

Soient x et y dans X , arbitraires.

$$\left(\sum_{i=1}^{n} F_i(y) - \sum_{i=1}^{n} F_i(x)\right)^T (y - x) = \sum_{i=1}^{n}(F_i(y) - F_i(x))^T(y - x)\ .$$

Comme F_i est co-coercive de constante α_{F_i} $\forall\ i \in \{1, \cdots, n\}$,

$$\begin{aligned}\sum_{i=1}^{n}\left[(F_i(y) - F_i(x))^T(y - x)\right] \ &\geq\ \sum_{i=1}^{n}\alpha_{F_i}\ \|F_i(y) - F_i(x)\|^2 \\ &\geq\ \sum_{i=1}^{n}\alpha_{F}\ \|F_i(y) - F_i(x)\|^2\end{aligned}$$

avec $\alpha_F = \min\limits_{i\in\{1,\cdots,n\}} \alpha_{F_i}$

$$\begin{aligned}&\geq\ \alpha_F \sum_{i=1}^{n}\ \|F_i(y) - F_i(x)\|^2 \\ &\geq\ \alpha_F\ \left\|\sum_{i=1}^{n} F_i(y) - \sum_{i=1}^{n} F_i(x)\right\|^2\end{aligned}$$

■

Références

[1] G. AUCHMUTY, Variational principles for variational inequalities, *Numerical Functional Analysis and Optimization*, Vol. 10, pp. 823–874, 1989.

[2] M.S. BAZARRA & C.M. SHETTY, *Nonlinear Programming: Theory and Algorithms*, John Wiley & Sons, New York, NY, 1979.

[3] C. BERGE, *Topological Spaces*, Oliver & Boyd, Edinburg, 1963.

[4] F.H. CLARKE, Generalized gradients and applications, *Transactions of the American Mathematical Society*, 205 (1975), pp. 247–262.

[5] F.H. CLARKE, A new approach to Lagrange multipliers, *Mathematics of Operations Research*, 1 (1976), pp. 165–174.

[6] G. COHEN, Auxiliary problem principle extended to variational inequalities, *Journal of Optimization Theory and Applications*, Vol. 59, pp. 325–333, 1988.

[7] I. EKELAND & R. TEMAM, *Convex Analysis and Variational Problems*, North-Holland, Amsterdam, 1976.

[8] W.W. HOGAN, Point-to-set maps in mathematical programming, *SIAM Review*, Vol. 15 (1973), pp. 591–603.

[9] E.S. LEVITIN & B.T. POLYAK, Constrained minimization methods, *USSR Computational Mathematics and Mathematical Physics*, 6 (1966), pp. 1–50.

[10] H. MINE and M. FUKUSHIMA, A minimization method for the sum of a convex function and a continuously differentiable function, *Journal of Optimization Theory and Applications*, 33 (1981), pp. 9-23.

[11] G.S. MINTY, Monotone (non linear) operators in Hilbert space, *Duke Mathematical Journal*, Vol. 29 (1962), pp. 341–346.

[12] A. MORDECAI, *Nonlinear Programming, Analysis and Methods*, Prentice-Hall, Inc., Englewood Cliffs, New Jersey, 1976.

[13] H. NIKAIDO, *Convex Structures and Economic Theory*, Academic Press, New York, NY, 1968.

[14] J.M. ORTEGA & W.C. RHEINBOLDT, *Iterative Solution of Nonlinear Equations in Several Variables*, Academic Press, New York, NY, 1970.

[15] M. PATRIKSSON, Partial linearization methods in nonlinear programming, Report LITHMAT-R-91-11, Department of Mathematics, Linköping Institute of Technology, Sweden, 1991. *Journal of Optimization Theory and Applications*, 78 (1993).

[16] M. PATRIKSSON, A unified description of iterative algorithms for traffic equilibria, Report LITHMAT-R-91-35, Department of Mathematics, Linköping Institute of Technology, Sweden, 1991. *European Journal of Operational Research* (1993).

[17] B.T. POLYAK, *Introduction to Optimization, Optimization Software*, New York, NY 1987.

[18] R.T. ROCKAFELLAR, Characterisation of the subdifferential of convex functions, *Pacific Journal of Mathematics*, 17 (1966), pp. 497–510.

[19] R.T. ROCKAFELLAR, *Convex Analysis*, Princeton University Press, Princeton, NJ, 1970.

[20] R.T. ROCKAFELLAR, On the maximal monotonicity of subdifferential mappings, *Pacific Journal of Mathematics*, 33 (1970), pp. 209–216.

[21] R.T. ROCKAFELLAR, On the maximality of sums of nonlinear monotone operators, *Transactions of the American Mathematical Society*, 149 (1970), pp. 555–562.

[22] R.T. ROCKAFELLAR, Monotone operators and the proximal point algorithm, *SIAM Control Optimization*, Vol. 14, No. 5, 1976.

[23] R.T. Rockafellar, *The Theory of Subgradients and its Applications to Problems of Optimization: Convex and Nonconvex Functions*, Heildermann Verlag, Berlin, 1981.

[24] F. Stoer and C. Witzgall, *Convexity and Optimization in Finite Dimensions, I*, Springer Verlag, Berlin, 1970.

[25] P. Tseng, Further applications of a splitting algorithm to decompsotion in variational inequalities and convex programming, *Mathematical Programming*, 48 (1990), 249–264.

[26] W.I. Zangwill, *Nonlinear Programming: A Unified Approach*, Prentice Hall, Englewood Cliffs, NJ, 1969.